STATISTICAL ANALYSIS PLAN

Protocol Number:

(099-37012)

SMARTPHONE ENABLED HEARING STUDY

Apple Inc.

[redacted]

By:

[redacted]

Biostatistics Lead

Initial Version Date: May 31, 2023
Current Version Number: [redacted]
Current Version Date: Aug 29, 2023

Table of Contents

Statistical Analysis Plan Approval Signature Page

The undersigned have reviewed and approve the Statistical Analysis Plan.

SIGNATURES

Sponsor Signatory	**Signature**	**Date**
Name/Role	_____	_____
[redacted]		
Name/Role	_____	_____
[redacted]		
Name/Role	_____	_____
[redacted]		
Name/Role	_____	_____
[redacted]		

All approvals are maintained in Agile.

LIST OF ABBREVIATIONS

4PTA	Pure Tone Audiometry (equivalent conduction audiometry average of four hearing level frequencies (0.5, 1, 2, and 4 kHz))
AE	Adverse Event
ACHIEVE	Aging and Cognitive Health Evaluation in Elders
CCAS	Complete Cases Analysis Set
CDC	Center for Disease Control
CV	Clinical Validation
dB	Decibel
FAS	Full Analysis Set
FIML	Full Information Maximum Likelihood
HL	Hearing Loss
Hz	Hertz
IOI-HA	International Outcome Inventory for Hearing Aids
kHz	Kilohertz
MAD	Mean Absolute Difference
MAR	Missing at Random
NAL	National Acoustic Laboratories
NL-2	NonLinear-2
OSPL90	Output Sound Pressure Level 90
PF	Pro-Fit
PI	Principle Investigator
PID	Patient Identifier
PPS	Per Protocol Analysis Set
PTA	Pure Tone Audiometry
REM	Real Ear Measure
SAS	Statistical Analysis Software
SAE	Serious Adverse Event
SIN	Speech In Noise
SF	Self-Fit
SPL	Sound Pressure Level
US	United States
UX	User Experience

1. [redacted]

Version	Summary of Changes	Author(s)/Role
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]

2. Introduction

This statistical analysis plan describes the analysis of the Clinical Validation (CV) portion of the Smartphone Enabled Hearing Study data.

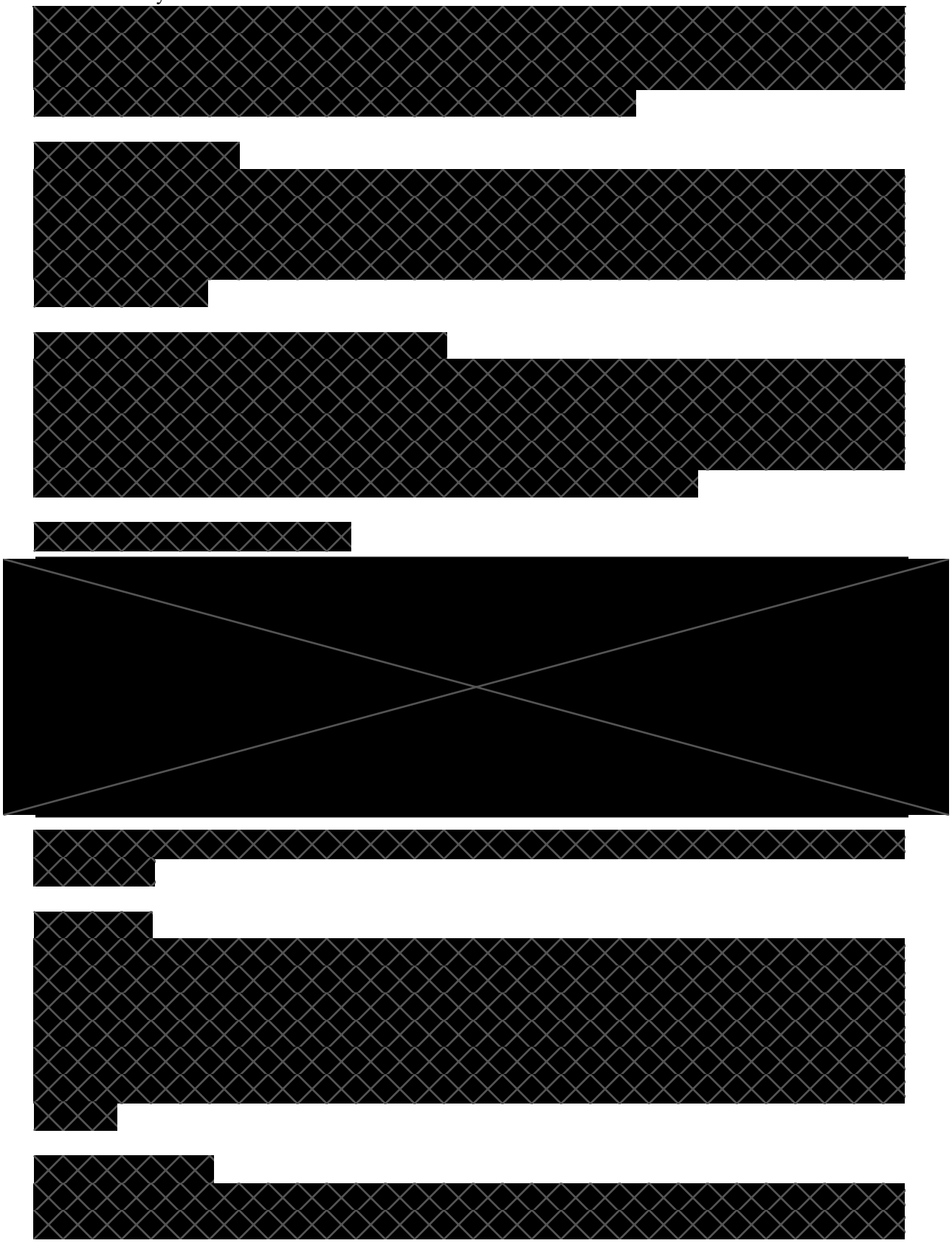
2.1. Background

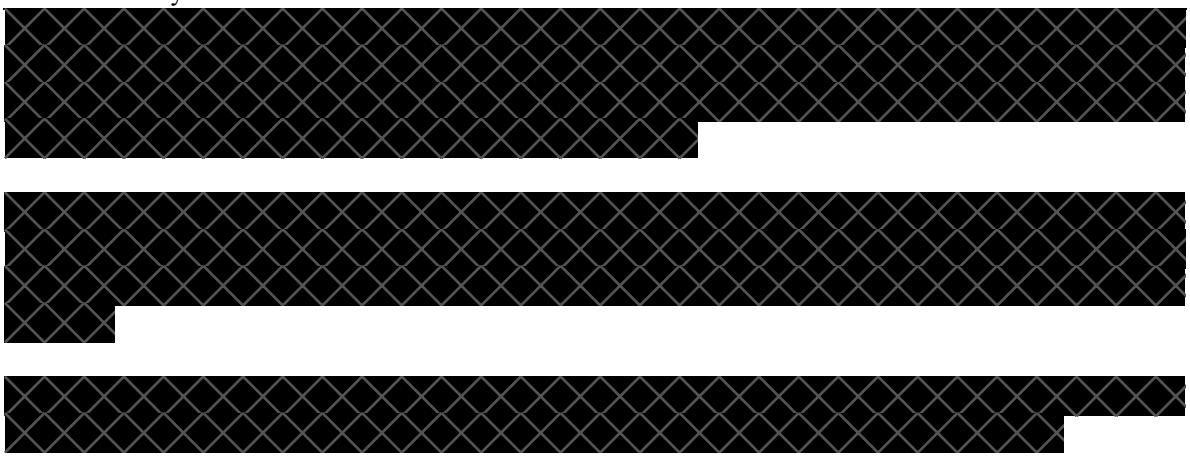
Hearing loss is the third most common chronic health condition, affecting nearly a quarter of Americans [4,5]. Age-related hearing loss is the most common cause and progresses over years from *mild* (20-34 dB hearing level [HL]) to, in increasing disability categories by 20 dB steps, *profound* (81 dB HL or greater) [1,4,5]. The majority of those affected have mild hearing loss, though individuals with normal range hearing, and those with unilateral hearing losses with normal hearing in their better ear, may also experience hearing problems [1,5]. Hearing loss is associated with significant comorbidities, including loneliness and social isolation [6]. Untreated hearing loss is associated with greater incident morbidity than no hearing loss across a range of health conditions, including dementia, depression, falls, and cardiovascular disease [7,8].

These comorbidities may occur many years before intervention for hearing loss, including in the mild to moderate hearing loss range where disability may be perceived as subtle [9]. This delay in intervention may be related in part to an access issue, as current standard is referral to an audiology clinic for testing and subsequent fitting for a hearing aid, which may occur over many months at significant cost in most regions of the United States [9]. This is related in-part to limited availability of audiologists and high volume of referrals. This supply and demand ratio is expected to negatively compound with the recent CDC (Centers for Disease Control and Prevention) recommendation for annual hearing screen for all adults with diabetes (11.3% of the US population) [10,11].

2.2. Device Description

The Hearing Aid Feature is a self-fitting air conduction Hearing Aid software for Over-the-Counter [redacted]





2.3. Study Rationale

Most who would benefit from aided hearing do not seek intervention, in part due to challenges with hearing aids after discovery. Hearing aid adoption rate is poor at approximately 33%, with a reported range of 10 or more years after initial diagnosis to time of first hearing aid fitting [9,12]. Reasons for this delay in care may include cost, stigma, and ease of use [9]. Delays in hearing aid may contribute to the comorbidities associated with the condition. The Aging and Cognitive Health Evaluation in Elders (ACHIEVE) trial was designed to assess for cognitive benefit in hearing loss intervention in an elderly population; findings are expected in 2023. In the pilot cohort, mental activity engagement increased in the hearing intervention group and decreased in the successful aging education intervention group that did not focus on improving hearing [13].

There is need for earlier and more efficient screening and intervention options with lower user barrier for adults with hearing problems. Here we propose a study to support decreased burden for hearing aid by smartphone enabled software settings for widely available wireless headphones.

3. Study Objectives

The primary objective of this study is to assess subjective non-inferiority in user-perceived benefit between a Self-Fitted (SF) (tuned) hearing assist settings group and a Professionally-Fitted (PF) (tuned) hearing assist settings group in participants with perceived or measured mild-to-moderate hearing loss.

Exploratory objectives are to assess objective improvement in hearing between a SF (tuned) hearing assist settings group and a PF (tuned) hearing assist settings group in participants with perceived or measured mild-to-moderate hearing loss.

4. Study Design Overview

This is a prospective intervention-based study that will involve data collection from individuals with measured or perceived mild-to-moderate hearing loss. This study is deemed to meet the qualifications of a research study with non-significant risk.

This study is designed to assess subjective improvement in hearing perceived by participants with mild- to moderate- hearing loss between an SF (tuned) hearing assist settings group and industry standard NAL-NL2 PF (tuned) hearing assist settings group [2]. An additional objective is to assess objective improvement in hearing between groups.

The study includes in-clinic data collection, and data collected in a participant's home environment or when completing daily activities.

After a screening visit, the study will involve at least three clinic visits to obtain a smartphone and wireless headphones, in-clinic assessment and measures, survey data, and two home intervals for optimization of hearing assist settings.

Preceding the start of the clinical validation study, a smaller group of participants will be enrolled into a sub-study [redacted]. These participants will follow similar study procedures outlined in this protocol. After the completion of the sub-study, the clinical validation study will commence at a later date. This sub-study is intended to provide assurance that all of the data collection, evaluation, and data transfer tools are performing as intended prior to the start of the clinical validation study or reveal unanticipated aspects of the study. As such, these participants are a separate cohort and will not be included in any of the endpoint analyses for the pivotal study.

As described in the protocol, approximately 112 participants will be enrolled in the clinical validation study to achieve minimum 90 useable data sets. The following details outline the demographic minimum recruitment targets for the clinical validation study population:

- At least 37 enrolled participants with female sex assigned at birth
- At least 45 enrolled participants should be <60 years of age

The minimum recruitment targets for categories of hearing loss are based upon the four pure-tone average (4PTA), which is the average of hearing levels measured at of 0.5, 1, 2, and 4 kHz, as measured in the ear with least hearing loss. The minimum recruitment targets for categories of hearing loss are as follows:

- Approximately 12 enrolled participants with No Impairment (4PTA: 15-25 dB HL) and perceived hearing loss where degree of hearing impairment is determined by the ear with least hearing loss.
- Approximately 37 enrolled participants with Mild hearing loss (4PTA: 26-40 dB HL) where degree of hearing impairment is determined by the ear with least hearing loss.

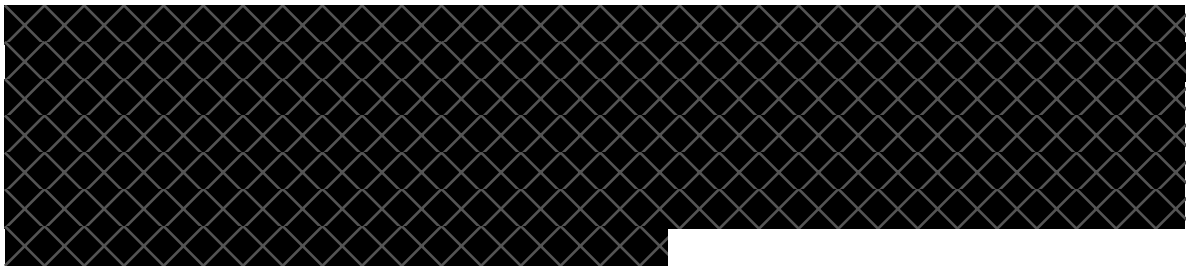
- Approximately 37 enrolled participants with Moderate hearing loss (4PTA: 41-60 dB HL) where degree of hearing impairment is determined by the ear with least hearing loss.

After a particular hearing loss recruitment category has been filled, subjects identified at screening within that category may be dispositioned prior to randomization. Information collected at the screening visit and clinic visit 1 (e.g., informed consent, medical history, demographics, otoscopy exam, perceived hearing loss questionnaire, eligibility criteria, and pure tone audiometry, tympanometry, and bone conduction) for these subjects will be presented in the listings but will not be included in any of the full analysis set analyses because they were not randomized.

5. Study Procedures

Participants will have a screening visit and up to three scheduled clinic visits interspersed by two home intervals. The participants enrolled into the sub-study cohort will follow similar study procedures:

• Screening Visit (~2 hours): informed consent, potential Pure Tone Audiometry (air conduction audiometry), medical history, and brief physical exam of the ears conducted by site clinician; time from Screening Visit to Clinical Visit 1 will be no longer than 3 weeks



• Clinic Visit 1 (~1.5 hour): examination of the ears by an audiologist who may remove ear wax if deemed required and consented to by participants, bone conduction audiometry, tympanometry, Pure Tone Audiometry (air conduction audiometry), enrollment, and baseline hearing assist software settings

• Home Interval 1 (3-8 days; begins on same day as Clinic Visit 1): at-home observations of baseline hearing assist software settings during participant's daily activities. Participants have the option to make setting adjustments during this interval for listening needs and/or to avoid discomfort, regular charging of study devices, and encouraged daily use of devices. Participants who have no evidence of at least daily use for several days will be contacted by study staff to assess for troubleshooting needs. Minimum expected wear time is at least 30 minutes per day for at least 3 days during this interval.

• Clinic Visit 2 (~1.5 hours): examination of the ears by an audiologist who may remove ear wax if deemed required and consented to by participants, hearing assist software setting fine-tuning, Real Ear measures, and SIN objective hearing test

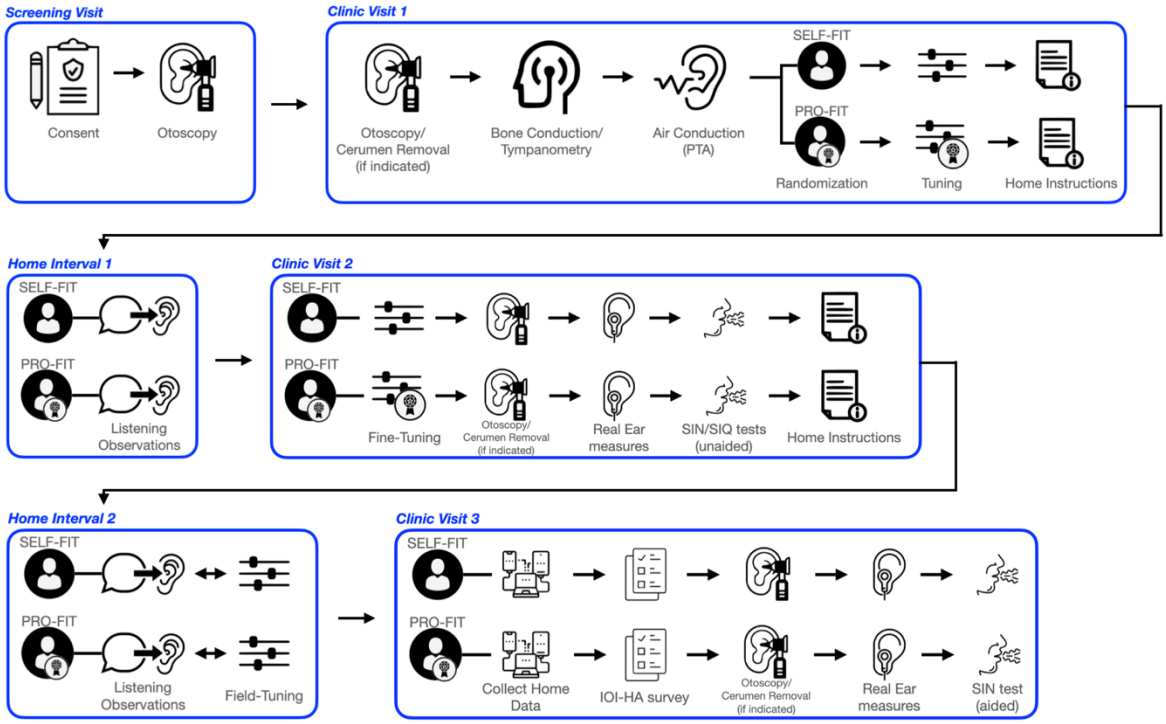
• Home Interval 2 (14-22 days; begins on same day as Clinic Visit 2): at-home fine-tuning of hearing aid settings during participant's daily activities, regular charging of study devices, and encouraged daily use of devices. Participants who have no evidence of at least daily use for several days will be contacted by study staff to assess for troubleshooting needs. Minimum expected wear time is at least 30 minutes per day for at least 14 days during this interval.

• Clinic Visit 3 (~1-2 hours): examination of the ears by an audiologist who may remove ear wax if deemed required and consented to by participants, IOI-HA survey[3], Real Ear Measures, SIN objective hearing test, and end of study items

• Participants may be asked to complete additional study procedures due to unforeseen circumstances such as replacing study equipment due to loss, theft, or damage.

Figure 5.1 depicts a diagram of study assessments and procedures.

Figure 5.1: Diagram of Study Assessments and Procedures



6. Study Endpoints

6.1. Primary Study Endpoints

The primary endpoint for the study is:

- IOI-HA Clinic Visit 3 scores. The average total IOI-HA scores will be compared between the Self-Fit Group and Pro-Fit Group

The International Outcome Inventory for Hearing Aids (IOI-HA) is a seven-item validated questionnaire designed to be generally applicable in evaluating the effectiveness of hearing aid treatments. The inventory was developed to facilitate cooperation among researchers and program evaluators in diverse settings. The survey has seven items assessing (1) daily hearing-aid use, (2) benefit, (3) residual activity limitation, (4) satisfaction, (5) residual participant restriction, (6) impact (of hearing impairment) on others, and (7) quality of life. Each of the 7 items are scored on a 1 (poorest) – 5 (best) Likert scale and summed to obtain a total outcome score ranging from 5 to 35 where higher scores are better.

6.2. Secondary Endpoints

There are no secondary endpoints in this study. Additional descriptive analyses will include the following:

- Comparison of Speech in Noise (SIN) scores between Self-Fit Group and Pro-Fit Group
- Comparison of Real Ear Measures (REM) between Self-Fit Group and Pro-Fit Group
- Summary of baseline pure tone audiometry (PTA) results
- Summary of baseline bone conduction audiometry results
- Summary of baseline tympanometry results

6.3. Study Hypotheses

The primary endpoint hypothesis for the clinical validation study is as follows:

H_0: Mean(IOI-$HA_{Pro\text{-}Fit}$) - Mean(IOI-$HA_{Self\text{-}Fit}$) > 3
H_1: Mean(IOI-$HA_{Pro\text{-}Fit}$) - Mean(IOI-$HA_{Self\text{-}Fit}$) $\leq$ 3

where Mean(IOI-$HA_{Pro\text{-}Fit}$) and Mean(IOI-$HA_{Self\text{-}Fit}$) are the average follow-up (i.e., clinical visit 3) IOI-HA scores for the Pro-Fit and Self-Fit groups, respectively.

The non-inferiority margin of 3 points was determined based on the estimated variation from mean scores previously reported (i.e., mean IOI-HA of 25.5 - 28.9) and is justified because 3 points is within a reasonable margin of variation for the study instrument at the subject level.

7. Statistical Considerations

7.1. Sample Size

Preceding the start of the clinical validation study, approximately [redacted] participants will be enrolled into a sub-study cohort. These participants will follow the same study procedures outlined in this protocol. After the completion of the sub-study, the clinical validation study

will commence. [redacted]

The approach to sample size determination of the clinical validation study is based on a two-sample independent t-test comparing group means assuming equal variances between groups. [redacted] a total of N=112 (56 per group) will be enrolled in the pivotal portion of the study. To summarize:

Clinical Validation Study: Approximately 112 participants will be enrolled

Sub-Study: Approximately [redacted] participants will be enrolled

7.2. Randomization and Blinding

Eligible subjects will be randomized (1:1) to either the Self-Fit or the Pro-Fit groups using a single pre-determined block randomization schedule with a fixed block size of 4. The randomization will not incorporate any pre-stratification. Subjects will not be blinded to their randomly assigned group.

7.3. Significance Level

The primary hypothesis test will use a one-sided significance level of 0.025. Analyses of the exploratory objectives will use 95% confidence intervals. Multiple comparison corrections will not be used.

7.4. Missing Data

Rigorous efforts will be made to ensure all subjects are compliant with the protocol. However, some subjects may drop out prematurely or some planned measurements may not be analyzable due to missing data. The handling of missing data for the primary endpoint analyses is described in Section 9.5. Exploratory efficacy analyses will be performed on available data. For subjects with one unanswered question on the IOI-HA, the mean of the other six questions will calculated to replace the missing data; otherwise, the IOI-HA survey score will be considered missing.

7.5. Subgroup Analyses

There are no planned subgroup analyses.

7.6. Interim Analyses

There are no interim analyses planned in this study.

8. Analysis Sets

The following analysis sets are defined for this study.

Eligible Analysis Set: All subjects who sign informed consent and meet eligibility criteria for the CV portion of the study. Some of these subjects may not have been randomized if the hearing loss recruitment category is already filled.

Full Analysis Set (FAS): All subjects who sign informed consent, meet eligibility criteria, and are enrolled and randomized into the CV portion of the study. This analysis set will be used for the primary endpoint analysis and to summarize subject accountability, protocol deviations, demographic and baseline characteristics, and adverse event data. Subjects in this analysis set will be analyzed "as randomized".

Per Protocol Analysis Set (PPS): All subjects who meet predetermined minimum data usability criteria, and who do not have any major protocol deviations that would necessitate exclusion from analysis. Subjects in this analysis set will be analyzed "as treated". This analysis set will be used as a supportive analysis of the primary endpoint. Subjects in the PP analysis set must meet the following two minimum usability criteria:

1) >10 minutes of Airpod wear for at least 1 day during home interval 1 AND
2) >10 minutes/day of Airpod wear for at least 7 days during home interval 2

The minimum wear times are justified to obtain the most accurate assessment of hearing aid benefit as measured by the IOI-HA questionnaire.

Complete Cases Analysis Set (CCAS): All subjects who have complete data for the efficacy endpoints of IOI-HA, QuickSIN, or Real Ear Measures. This analysis set will be used as a supportive analysis of the primary endpoint.

9. Analysis Approach

9.1.Subject Accountability

A summary of the number of eligible subjects, randomized subjects, withdrawn subjects, and completers will be presented by randomized treatment group (if applicable) and overall. Reasons for withdrawal will be summarized according to the following categories:

- Adverse Event
- Loss of Study Provided Materials
- Protocol Violation
- Does Not Meet Inclusion Exclusion Criteria After Enrollment

- Withdrawal of Consent by Subject
- PI Discretion
- Termination of Study
- Lost to Follow-up
- Other

Results of the perceived hearing loss questionnaire will also be summarized descriptively.

9.2. Device Measurement Accountability

A summary table will be reported which presents total duration of device wear time (hours) across home intervals 1 and 2 and the percentage of days worn where the audio settings were changed. These results will be presented overall and by randomized treatment group. Device dispensing and return information will be presented in the listings.

9.3. Demographic Characteristics

Descriptive statistics (e.g., N, Mean, Std. Dev., Min, Max) for continuous data types and frequencies for categorical data types will be displayed for the demographic characteristics by randomized treatment assignment and overall. The demographic characteristics and data types are listed below:

Table 9.1: Demographic Characteristics

Characteristic	Data Type
Age at Enrollment (years)	Continuous
Age Group at Enrollment	Categorical
Sex	Categorical
Ethnicity	Categorical
Race	Categorical

The following age group categories will be used: <60, and ≥60 years. Subjects may choose more than one Race category.

9.4. Baseline Characteristics

The following baseline characteristics will be summarized using descriptive statistics by randomized treatment group and overall.

9.4.1. Otoscopy Examination

Otoscopy results at screening will be summarized with frequencies and percentages separately for the right and left ear canals according to the following clinical categories:

- Normal
- Blocked, Collapsed
- Abnormal-Not Clinically Significant

- Abnormal-Clinically Significant
- Other

9.4.2. Bone Conduction Audiometry and Tympanometry Results

The bone conduction audiometry results (dB HL) at clinic visit 1 will be summarized descriptively (N, Mean, Std. Dev., Median, Min-Max) separately for each ear at each of the following four frequencies: 0.5kHz, 1 kHz, 2 kHz, and 4kHz.

Tympanometry results at clinic visit 1 will be summarized descriptively (frequencies and percentages) separately for each ear by the following response categories:

- Type A
- Type B
- Type C
- Other

9.4.3. Pure Tone (Air Conduction) Audiometry Results

The pure tone (air conduction) audiometry results (dB HL) at clinic visit 1 will be summarized descriptively (N, Mean, Std. Dev., Median, Min-Max) separately for each ear at each of the following eight frequencies: 250Hz, 500Hz, 1000Hz, 2000Hz, 3000 Hz, 4000Hz, 6000Hz, and 8000Hz. Additionally, the four tone average PTA (averaged across the 500, 1000, 2000, and 4000Hz frequencies separately for each ear) will be computed for each subject and rounded to the nearest whole integer. The ear with the least (i.e., lowest) hearing loss will be used for categorizing subjects into the following categories:

- 15 – 25 dB HL (No Impairment)
- 26 – 40 dB HL (Mild Impairment)
- 41 – 60 dB HL (Moderate Impairment)

9.4.4. Medical History

Relevant medical history information (i.e., Ear/Nose/Throat; Dementia or other neurological condition preventing following instructions; Vascular) will be coded using MedDRA v24.1 and summarized by preferred term.

9.5. Primary Endpoint Analysis

The pivotal primary endpoint analysis will be performed on subjects from the Full Analysis Set (FAS) with supportive analyses also conducted on the Per Protocol (PP) and Complete Cases (CC) Analysis Sets as described below.

Missing data for the IOI-HA score at clinic visit 3 will be assumed to be missing at random (MAR). Under this assumption, the probability that the IOI-HA is missing depends only on available information (e.g., baseline demographic and other baseline characteristics). As such, the first step in modeling the missing data is to perform a logistic regression where the

outcome variable is 0 for observed cases and 1 for missing cases. This logistic regression model will include the following information available at baseline: treatment group, age, sex, race (white vs. not white), pure tone audiometry results (average across the 500, 1000, 2000, and 4000 Hz frequencies for "best" ear (i.e., lowest value of the two ears)), and otoscopy and tympanometry exam results. Factors in this model will be considered significantly related to the probability of missing if the p-value is less than 0.15. PROC LOGISTIC in SAS will be used for this analysis. The otoscopy results will be dichotomized as "normal" vs. "not normal" for the best ear. Likewise, the tympanometry exam results will be dichotomized as "type A" vs. "not type A" for the best ear. Factors which confound with one another may be removed from the model to ensure model stability and convergence.

Once the set of significant factors has been identified via logistic regression, a full information maximum likelihood (FIML) analysis will be performed using PROC CALIS in SAS to estimate the treatment effect after adjustment for significant baseline information. This approach has been shown to produce unbiased parameter estimates and standard errors under MAR [17]. This approach also provides a unique estimate of the treatment difference (PF – SF) and requires fewer decisions than multiple imputation. If there are no baseline factors identified through the logistic regression analysis which are significantly related to missingness and if there is no missing data for any of the baseline factors, the parameter estimate of the treatment effect will be the arithmetic mean difference (PF – SF) between the two groups but with a smaller standard error than using only complete cases.

The process works by estimating a likelihood function for each study subject based on the baseline factors that are present so that all the available data are used. Model fit information is derived from a summation across fit functions for individual subjects, and, thus, model fit information is based on all subjects. With missing data, the FIML fit function [18] is computed for each set of subjects with the same unique pattern of missing values—a casewise likelihood. So, an i subscript is used in the equation below to show that the fit function is for each particular subject:

$$\log L_1 = -\frac{k_i}{2}\log(2\pi) - \frac{1}{2}\log\left|\boldsymbol{\Sigma}_i\right| - \frac{1}{2}\left(\mathbf{Y}_i - \boldsymbol{\mu}_i\right)^T \boldsymbol{\Sigma}_i^{-1}\left(\mathbf{Y}_i - \boldsymbol{\mu}_i\right)$$

where log is the natural logarithm (with base e), π is the mathematical constant, Σ_i is the population covariance matrix, $\mathbf{Y}$ is vector of observed variables, μ_i is the vector of their means, and the superscript T is the matrix transpose function. The subscripts i denote individual subjects in which the $\mathbf{Y}_i$ vectors differ in length depending on the number of non-missing data values. Similarly, matrices Σ_i and μ_i vary by deleting rows and columns for missing variables. The sum of the individual log likelihoods is then computed. This process approach allows one to use all the available information in the variables. Additionally, as can be seen from this likelihood function, this approach assumes the IOI-HA scores follow a normal distribution.

Because the study hypothesis is a test of non-inferiority, a Wald-type 95% two-sided confidence interval for the treatment effect will be computed (i.e., treatment effect parameter

estimate +/- 1.96*(Standard error of the estimated treatment effect)) and compared to the non-inferiority margin of 3. If the upper 95% confidence bound for the treatment effect (defined such that higher values favor the pro-fit group) is less than or equal to 3, the null hypothesis will be rejected and the self-fit treatment arm will be considered non-inferior to that of the pro-fit treatment arm.

The pivotal primary endpoint analysis using the above methodology will be based on the Full Analysis Set where all randomized subjects are included in the analysis and subjects will be analyzed according to the treatment group to which they were randomized. A supportive analysis will also be performed using the Per Protocol Analysis Set where subjects will be analyzed according to the treatment received.

A Complete Cases analysis using only those subjects who have a non-missing outcome for the IOI-HA survey will also be performed using a two-sample independent non-inferiority t-test assuming equal variances. The two-sided 95% confidence interval will be reported for the (Pro-Fit – Self-Fit) mean difference. In the event there are substantial departures from the equal variances assumption, a two-sample t-test assuming unequal variances will be performed similarly.

If the non-inferiority hypothesis is rejected using the Full Analysis Set, a gatekeep strategy will be used to test for superiority to determine if the treatment effect (i.e., mean difference) is statistically different than 0 using the same two-sided significance level of 0.05. If the upper two-sided 95% confidence bound for the treatment effect (Pro-Fit – Self-Fit) is less than or equal to 0, superiority of the Self-Fit arm will also be concluded in addition to non-inferiority.

9.6. Secondary Endpoint Analyses

There are no secondary endpoints to be analyzed.

9.7. Additional Analyses

Three different Speech in Noise (SIN) test lists will be administered at clinic visit 2 (unaided) and again at clinic visit 3 (aided). The three lists were specifically chosen based on their homogeneity and as such, average test scores will be computed as the average of the test scores from the three lists. Average test scores and test scores from each of the three lists will be summarized descriptively for each assessment visit for each of the two randomized groups (Self-Fit and Pro-Fit) along with corresponding 95% confidence intervals for the means based on the t-distribution. Additionally, the mean difference between treatment groups of the within-subject differences (clinic visit 3 – clinic visit 2) will be reported along with the corresponding two-sided 95% confidence interval.

For the real ear measures (REM) testing, the REM output values (dB SPL) will be summarized descriptively from all available ears at each frequency (0.25, 0.5, 1, 2, 3, 4, 6, and 8 kHz) and averaged across all frequencies separately for each of the two study arms (Self-Fit and Pro-Fit) for each of the two clinic visits. These REM analyses will be

performed separately at 50, 65, and 80 dB SPL. The mean absolute difference (MAD) between the Self-Fit and NAL-NL2 targets and the MAD between the Pro-Fit and NAL-NL2 targets will also be computed for each clinic visit at each frequency (0.5, 1, 2, and 4 kHz) and averaged across all frequencies using all available ears. Two-sided 95% bootstrap confidence intervals (sampling subjects with replacement) for the MAD values will also be reported.

9.8. Safety Analyses

All adverse events will be recorded throughout the entire study period, whether they are considered to be related to the study procedures or not. Signs and symptoms of each AE will be described in detail: date of event, description of event, severity, relationship to study procedures, action taken and outcome. Adverse events will be collected as spontaneously reported by the subjects.

Adverse events will be coded using MedDRA v24.1. The number of any adverse events and the number and percentage of subjects reporting each type of adverse event will be presented by preferred term. Multiple occurrences of the same event reported by the same subject will be counted only once.

Adverse event summaries (number of events and incidence) will be presented for:

- All adverse events
- Serious adverse events (SAEs)
- Study procedure related adverse events
- Severe adverse events

The intensity of an AE will be categorized as follows:

- Mild: Events that are transient, easily tolerated by the participant, and does not affect the participant's daily activities
- Moderate: Events that cause the participant discomfort and interrupts the participant's usual daily activities. May require treatment but not extended hospitalization for the participant.
- Severe: Events that are incapacitating, causing inability to do work or usual activities; signs and symptoms may require medical evaluation and/or treatment, and may require additional prolonged hospitalization

Causal relationship assessment to study procedures is required for purposes of reporting AEs. To promote consistency, the following guidelines should be taken into consideration along with good clinical and scientific judgment when determining the relationship of study procedures to an AE:

- Probably Related: There is a strong temporal relationship to the study or device, and an alternative etiology for the AE is unlikely.

- Possibly Related: The event is not readily explained by the participant's medical condition, concomitant therapy, or other causes, and there is a plausible temporal relationship between the event and participation in study procedures.

- Probably Not Related: An event, for which an alternative explanation is more likely (e.g., concomitant medications or ongoing medical conditions) or the temporal relationship to study procedures and/or exposure suggests that a causal relationship is unlikely.
- Not Related: An AE that does not follow a reasonable temporal sequence from participating in study procedures and/or that can reasonably be explained by other factors, such as underlying diseases, complications, concomitant drugs, and concurrent treatments.

Study procedure related events will include those events classified as possibly or probably related to the study procedure.

Action taken will be defined as:
- None;
- Study procedures interrupted;
- Study procedures stopped

Outcome will be defined as:
- Resolved;
- Ongoing or stabilized and followed by private healthcare provider;
- Lost to follow up

9.9. Medication Usage

No prior or concomitant medication usage will be recorded in this study other than the following two exclusionary criteria:
- Active treatment, or treatment in the past 6 months, with either a chemotherapeutic drug for cancer, or radiation therapy to the head or neck region
- Active treatment, or treatment in the past 6 months, with parenteral aminoglycoside antibiotics

9.10. Protocol Deviations

Protocol deviations will be summarized descriptively by deviation type and deviation severity for each randomized treatment group and overall.

10. Statistical Software

All analyses will be performed with SAS (v9.4 or later) and/or R (v3.6.0 or later).

11. Changes to Planned Analysis

All changes to planned analyses will be documented in the clinical study report.

12. References

1. Olusanya BO, Davis AC, Hoffman HJ. Hearing loss grades and the International classification of functioning, disability and health. Bull World Health Organ 2019; 97(10):725-728
2. Keidser G, Dillon H, Flax M, et al. The NAL-NL2 prescription procedure. Audiol Res 2011; 1(1):e24
3. Hearing Aid Research Lab (HARL) (last checked October 2022). International Outcome Inventory for Hearing Aids (IOI-HA). https://harlmemphis.org/international-outcome-inventory-for-hearing-aids-ioi-ha/
4. Nieman CL, Reed NS, Lin FR. Otolaryngology for the Internist: Hearing Loss. Med Clin North Am 2018; 102(6):977-992
5. Goman AM, Lin FR. Prevalence of hearing loss by severity in the United States. Am J Public Health 2016;106(10):1820-1822
6. Shukla A, Harper M, Pedersen E, et al. Hearing Loss, Loneliness, and Social Isolation: A Systematic Review. Otolaryngol Head Neck Surg 2020;162(5):622-633
7. Lin FR, Metter EJ, O'Brien RJ, et al. Hearing loss and incident dementia. Arch Neurol 2011;68(2):214-220
8. Deal JA, Reed NS, Kravetz AD, et al. Incident Hearing Loss and Comorbidity: A Longitudinal Administrative Claims Study. JAMA Otolaryngol Head Neck Surg 2019;145(1):36-43
9. Grundfast KM, Liu SW. What Otolaryngologists Need to Know About Hearing Aids. JAMA Otolaryngol Head Neck Surg 2017;143(2):109-110
10. Centers for Disease Control and Prevention (2022, May 27). Prevent Diabetes Complications: Diabetes and Hearing Loss. https://www.cdc.gov/diabetes/managing/diabetes-hearing-loss.html
11. Centers for Disease Control and Prevention (2022, Jan 18). National Diabetes Statistics Report: Estimates of Diabetes and Its Burden in the United States. https://www.cdc.gov/diabetes/data/statistics-report/index.html
12. Davis A, Smith P, Ferguson M, et al. Acceptability, benefit and costs of early screening for hearing disability: a study of potential screening tests and models. Health Technol Assess 2007;11(42):1-294
13. Goman AM, Gao T, Betz J, et al. Association of Hearing Loss with Physical, Social, and Mental Activity Engagement. Semin Hear 2021;42(1):59-65
14. Saunders GH, (2005, August – 2008, November). The performance-perceptual test as a counseling tool (PPT). Identifier NCT00126113. https://clinicaltrials.gov/ct2/show/study/NCT00126113?term=international+outcomes+inventory+hearing+aids&draw=3&rank=12
15. Smith S & VA Office of Research and Development, (2006, November - 2008, October). Predicting success with hearing aids. Identifier NCT00371449. https://clinicaltrials.gov/ct2/show/record/NCT00371449?term=international+outcomes+inventory+hearing+aids&draw=3&rank=13
16. Houmøller SS, Wolff A, Möller S, et al. Prediction of successful hearing aid treatment in first-time and experienced hearing aid users: Using the International Outcome Inventory for Hearing Aids. International Journal of Audiology 2022;61(2):119-129

17. Enders, C. K., & Bandalos, D. L. The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling,* 2001;8(3), 430–457. https://doi.org/10.1207/S15328007SEM0803_5
18. Arbuckle JL. Full information estimation in the presence of incomplete data. In G.A. Marcoulides and R.E. Schumacker [Eds.] Advanced structural equation modeling: Issues and Techniques (243-277). 1996. Mahwah, NJ: Lawrence Erlbaum Associates.

13. List of Tables/Listings/Figures

List of Tables

List of Listings

List of Figures

Figure 3.1.1	Mean Absolute Difference (MAD) of Real Ear Measures (REM) from NAL-NL2 Targets Per Frequency by Treatment Group at Clinic Visit 2 at 50 dB SPL– Complete Cases Analysis Set
Figure 3.1.2	Mean Absolute Difference (MAD) of Real Ear Measures (REM) from NAL-NL2 Targets Per Frequency by Treatment Group at Clinic Visit 2 at 65 dB SPL– Complete Cases Analysis Set
Figure 3.1.3	Mean Absolute Difference (MAD) of Real Ear Measures (REM) from NAL-NL2 Targets Per Frequency by Treatment Group at Clinic Visit 2 at 80 dB SPL– Complete Cases Analysis Set
Figure 3.2.1	Mean Absolute Difference (MAD) of Real Ear Measures (REM) from NAL-NL2 Targets Per Frequency by Treatment Group at Clinic Visit 3 at 50 dB SPL- Complete Cases Analysis Set
Figure 3.2.2	Mean Absolute Difference (MAD) of Real Ear Measures (REM) from NAL-NL2 Targets Per Frequency by Treatment Group at Clinic Visit 3 at 65 dB SPL- Complete Cases Analysis Set
Figure 3.2.3	Mean Absolute Difference (MAD) of Real Ear Measures (REM) from NAL-NL2 Targets Per Frequency by Treatment Group at Clinic Visit 3 at 80 dB SPL- Complete Cases Analysis Set
Figure 4.1	Mean Absolute Difference (MAD) of Real Ear Measures (REM) from NAL-NL2 Targets Averaged Across all Frequencies by Treatment Group at Clinic Visit 2– Complete Cases Analysis Set
Figure 4.2	Mean Absolute Difference (MAD) of Real Ear Measures (REM) from NAL-NL2 Targets Averaged Across all Frequencies by Treatment Group at Clinic Visit 3– Complete Cases Analysis Set