Statistical Analysis Plan

Evaluating the clinical effectiveness of a community-based hearing aid fitting service delivery model facilitated by CHWs providing smartphone-based in-situ and pre-set hearing aid fittings in low- and middle-income communities

Version 1.0
Trial Registration Number: NCT06982716
Date: 14 August 2025

Authors: Karina C. De Sousa, Caitlin Frisby, De Wet Swanepoel, Marien A. Graham, David R. Moore, Lisa Hunter

Corresponding author:

De Wet Swanepoel
dewet.swanepoel@up.ac.za

The University of Pretoria
Department of Speech-Language Pathology and Audiology

Contents

1. Administrative Information

1.1. Study identifiers

- Research Ethics Approval- University of Pretoria, Health Sciences Research Ethics Committee Approval Number: HUM011/0724)
- Clinical trial registry- clinicaltrials.gov (Identifier: NCT06982716)

1.2. Contributors the statistical analysis plan

Name and ORCID ID:	Primary Affiliation	Role on the study	SAP contribution
Caitlin Frisby https://orcid.org/0000-0002-9922-2927	University of Pretoria, Department of Speech-Language Pathology and Audiology, South Africa	Research coordinator & Investigator	Prepared initial draft and statistical analyses
Karina C. De Sousa https://orcid.org/0000-0003-1742-1613	University of Pretoria, Department of Speech-Language Pathology and Audiology, South Africa	Research coordinator & Investigator	Prepared initial draft and statistical analyses
De Wet Swanepoel https://orcid.org/0000-0001-8313-1636	University of Pretoria, Department of Speech-Language Pathology and Audiology, South Africa	Co - Primary Investigator	Prepared initial draft and revised statistical analyses plan
David R. Moore https://orcid.org/0000-0002-1567-1945	Cincinnati Children's Hospital Medical Centre, Cincinnati, United States	Co-Primary Investigator	Reviewed draft
Lisa Hunter https://orcid.org/0000-0003-4057-567X	Cincinnati Children's Hospital Medical Centre, Cincinnati, United States	Co-Primary Investigator	Reviewed draft
Marien A. Graham https://orcid.org/0000-0003-4071-9864	University of South Africa, Department of Mathematics Education	Study statistician	Reviewed draft and critically revised statistical analyses plan

2. Study site and investigators

2.1. Study site

The clinical study will take place at three sites across two Provinces in South Africa.

Community centres or homes (if participants do not have access to the community centre) across:

- Atteridgeville District, Gauteng, South Africa
- Khayelitsha District, Western Cape, South Africa
- Drakenstein District, Western Cape, South Africa

2.2. Study investigators and administrative structure

Table 1 describes the roles and responsibilities assigned to clinical research team members.

Table 1. Study personnel

Role	Name	Responsibility
Co-Principal Investigator	De Wet Swanepoel	Study planning, oversight, data analyses, and reporting
Co-Principal Investigator	David R Moore	Study planning, oversight, data analyses, and reporting
Co-Principal Investigator	Lisa Hunter	Study planning, oversight, data analyses, and reporting
Co-investigator	Herman Myburgh	Study planning, technical oversight, data analyses, and reporting
Co-investigator & Statistician	Marien A Graham	Study planning, technical oversight, data analyses, and reporting
Site Coordinator	Tersia de Kock	Study planning, technical oversight, data collection, analysis, and reporting. Site coordinator for Khayelitsha and Drakenstein district
Research Coordinator	Karina De Sousa	Study planning, technical oversight, data collection, data analyses, and reporting
Research Coordinator & Site Coordinator	Caitlin Frisby	Study planning, technical oversight, data analyses, and reporting. Site coordination for Atteridgeville district

3. Introduction

3.1. Overview of the study

Hearing loss is a leading cause of disability worldwide, affecting over 1.5 billion people in 2020, with more than 430 million experiencing disabling hearing loss (World Health Organization, 2021). The burden is disproportionately concentrated in low- and middle-income countries (LMICs), where 90% of individuals with moderate to profound hearing loss reside, yet access to hearing health services is severely limited (Mulwafu et al., 2017; WHO, 2021). In sub-

Saharan Africa, hearing aid uptake is below 3% among those in need, highlighting a significant treatment gap (Bisgaard et al., 2021).

The World Health Organization has emphasized the use of digital health solutions and task-shifting strategies, such as deploying trained community health workers (CHWs), to improve hearing care access. Over-the-counter (OTC) and pre-set hearing aids, especially when combined with mobile health (mHealth) technologies, are increasingly recommended as scalable interventions suitable for LMICs (WHO, 2021; 2024).

This study evaluates a community-based service delivery model that leverages CHWs and mHealth-supported hearing aid fittings. The aim is to assess the effectiveness of two forms of hearing aid fitting (smartphone-based in-situ and pre-set amplification) when facilitated by CHWs in low-resource settings. A key feature of the design is the inclusion of a minimal amplification control group, providing a flat 10 dB gain across frequencies without individual tailoring. This control condition serves to isolate the therapeutic benefit of meaningful amplification from non-specific effects such as device wear or participant expectations.

The study also incorporates a standardized, paper-based acclimatization and support programme, provided to all participants at the time of fitting, as well as follow-up access to CHWs for additional support. The intervention is designed to address common barriers to hearing aid use in LMICs, such as limited follow-up, poor device management, and lack of user education.

The results of this trial aims to inform the potential scale-up of CHW-facilitated hearing aid delivery as a sustainable model of care. Furthermore, by directly comparing experimental interventions to minimal amplification, the study aims to determine whether self-reported outcomes can be attributed to true therapeutic benefit, thereby guiding future clinical and policy decisions in underserved populations.

3.2. Study design

3.2.1. General study design

This study is a randomized, three-arm, single-blind, placebo-controlled trial designed to evaluate the effectiveness of community-based hearing aid fittings facilitated by community health workers (CHWs) in low-resource settings. A total of 90 participants will be enrolled and followed for up to 52 weeks after initial hearing aid fitting.

The trial will be conducted across three sites in South Africa: Khayelitsha and Drakenstein Districts (Western Cape Province), and Atteridgeville District (Gauteng Province).

Participants will be randomly allocated in a 1:1:1 ratio to one of three intervention arms:

- Arm 1: In-situ fitting - Smartphone-based, personalized fitting using Lexie Lumen hearing aids programmed via a proprietary algorithm aligned with the NAL-NL2 prescription.
- Arm 2: Pre-set fitting - Go Ultra hearing aids fitted using a standard pre-set configuration without individualized adjustment.

- Arm 3: Minimal gain fitting (Control) - Lexie Lumen hearing aids set to provide a flat, non-individualized gain of 10 dB across frequencies, serving as a placebo condition. This group will be crossed over to the in-situ fitting after the 6-week follow-up.

Randomization will occur prior to enrollment, ensuring allocation concealment. Participants will be assigned to one of the three arms before undergoing any intervention. Full randomization procedures, including block size and stratification, are described in Section 3.2.2 of this SAP. This is a single-blind trial: participants are unaware of their group allocation. Outcome assessors are also blinded to allocation. However, due to the nature of the intervention, CHWs facilitating the fittings are not blinded.

Study Timeline

- T0 (Baseline): Screening and eligibility assessment.
- T1 (Fitting): Participants receive their assigned hearing aid configuration and begin a 6-week field trial.
- T2 (6 weeks): Primary endpoint assessment. At this point, the minimal gain group (control) will be crossed over to the in-situ fitting condition (CG-T2), and a second 6-week trial will commence for this group.
- T3 (12 weeks): Secondary endpoint assessment for In-situ and Pre-set groups; first follow-up for the control group post-crossover (CG-T3).
- T4 (26 weeks) and T5 (52 weeks): Long-term follow-up assessments for the experimental arms.

See Figure 1 for a schematic representation of the study timeline and crossover process.

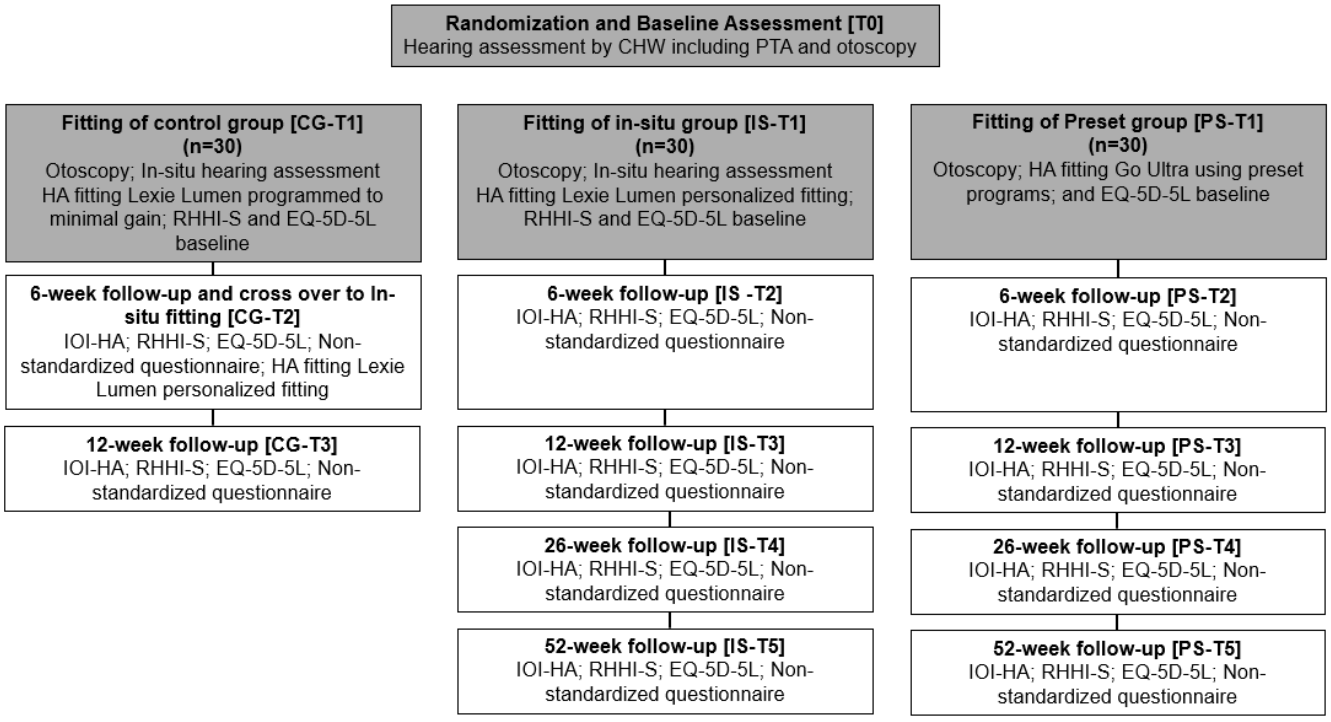

Figure 1. Proposed clinical trial design.

3.2.2. Randomization

Randomization will be implemented to minimize allocation bias and ensure balance across the three study arms: In-situ fitting, Pre-set fitting, and Minimal gain (control). A total of 90 participants (*n* = 30 per arm) will be enrolled across three study sites in South Africa. A permuted block randomization strategy with fixed block sizes of three (n = 3) will be used. This approach ensures equal distribution of participants across the study arms at regular intervals throughout the enrolment process, while reducing predictability in group assignment. Within each block, participants will be randomly allocated to one of the three intervention arms, with the order of allocation randomized to maintain allocation concealment. The randomization sequence will be computer-generated prior to trial initiation using a secure, web-based randomization tool such as Sealed Envelope (https://www.sealedenvelope.com/simple-randomiser/v1/lists) or a comparable platform. The final sequence will be stored in a password-protected, centralized Google Sheet, accessible only to designated site administrators for allocation purposes.

Allocation Concealment and Implementation

The allocation sequence will be prepared in advance by the trial statistician and will remain concealed from investigators and CHWs conducting participant recruitment and intervention delivery. Upon completion of eligibility screening, the site administrator will assign participants to their respective group using the next available allocation from the centralized sheet. This process ensures pre-enrolment randomization and maintains blinding of participants and outcome assessors to treatment allocation. This centralized, pre-defined allocation system is designed to preserve the integrity of the trial, prevent selection bias, and support reproducibility of the randomization process.

3.3. Study interventions

Table 2 below compares the interventions of the three groups of the study.

Table 2. Comparison of the three interventions

Feature/characteristic	**In-situ fitting (Lexie Lumen)**	**Pre-set Fitting (Go Ultra)**	**Minimal Gain Fitting (Lexie Lumen-Control)**
Device type	Self-fitting OTC hearing aid	Pre-set OTC hearing aid	Self-fitting OTC hearing aid
Device brand/model	Lexie Lumen	Go Ultra	Lexie Lumen
Fitting approach	In-situ audiometry via smartphone app (CHW-assisted)	Pre-defined gain settings (CHW-assisted)	Flat 10 dB gain across frequencies (CHW-assisted)
Programming method	Smartphone app using proprietary NAL-NL2 based fitting algorithm	Fixed internal pre-set programs	Smartphone app-manual override after in-situ test
Number of channels	16 channels	16 channels	16 channels
Listening programs	4 programs: Everyday, Noisy, Indoor, Outdoor, Music	4 user-selectable pre-set programs	4 programs: Everyday, Noisy, Indoor, Outdoor, Music
Noise reduction or directionality	Digital noise reduction, directional microphones	Digital noise and wind noise reduction	Same features as in-situ device
Battery type	Replaceable	Rechargeable	Replaceable
Memory Functions	Program and volume memory	Program and volume memory	Program and volume memory

Target population	Adults with mild to moderate hearing loss	Adults with mild to moderate hearing loss	Adults with mild to moderate hearing loss

3.4. Objectives and endpoints

The primary objective of this study is to evaluate self-reported hearing aid outcomes, measured using the IOI-HA global score at 6 weeks post-fitting, across three intervention arms. Two primary hypotheses will be tested: first, a superiority hypothesis, assessing whether hearing aids fitted using pre-set or in-situ self-fitting methods result in statistically and clinically significant improvements compared to hearing aids fitted with minimal gain; and second, a non-inferiority hypothesis, evaluating whether the pre-set fitting method is non-inferior to the in-situ fitting method. Both hypotheses will be analysed using the same primary outcome measure and time point. Additionally, an exploratory objective is to examine the broader impact of hearing aid use on an additional self-reported hearing outcome (RHHI-S) and health-related quality of life across all intervention groups.

The following outcome measures will be included in the trial comparison:

Table 3. Study endpoint measures	
Primary endpoint measure	
Endpoint	**Description**
International Outcome Inventory for Hearing Aids (IOI-HA)	The IOI-HA is a validated seven-item questionnaire to measure the effectiveness of the hearing aid intervention (Cox and Alexander, 2002). It targets seven domains including, (i) daily use, (ii) benefit, (iii) residual activity limitations, (iii) satisfaction, (iv) residual participation restrictions, (v) impact on others, and (vi) quality of life. Each item has five response choices, from worst to best outcome. The IOI-HA has been translated to the languages most commonly used in these communities namely isiXhosa and Sepedi. The participants will have the option to complete the IOI-HA in either of these languages or in English.
Exploratory variables	
Revised Hearing Handicap Inventory- Screening (RHHI-S)	The RHHI-S is a validated 10-item questionnaire that is a strong unidimensional, clinically informative measure of self-perceived hearing handicap that can be used by adults of all ages (Cassarly, Matthews, Simpson, & Dubno, 2020). Each question has three possible responses, including yes, sometimes, or no. The RHHI-S has been translated into the languages most commonly used in these communities, namely isiXhosa and Sepedi. The participants will have the option to complete the RHHI-S in either of these languages or in English.
EQ-5D-5L	The EQ-5D-5L is a standardized instrument developed by the EuroQol Group to assess health-related quality of life. It consists of a Descriptive System. This covers five dimensions—mobility, self-care, usual activities, pain/discomfort, and anxiety/depression—each with five levels of severity (from "no problems" to "extreme problems"). The respondent selects one level in each dimension, resulting in a 5-digit health state profile (e.g., 1-2-3-1-1) {Herdman, 2011 #9765}. Each 5-digit health state profile is converted into a single index value (utility score) using a country-specific value set. The utility score reflects the individual's overall health status, where 1 = full health, 0 = dead, <0 = states worse than death. The EQ-5D-5L has been translated into the

	languages most commonly used in these communities, namely isiXhosa and Northern Sotho (Sepedi). The participants will have the option to complete the EQ-5D-5L in either of these languages or in English.
Berkman-Syme Social Network Index (SNI)	The Berkman-Syme Social Network Index (SNI) is a well-established tool used to assess the extent of an individual's social connections. The SNI examines the relationship between social networks and health outcomes. The SNI evaluates social integration by assessing four key domains: Marital status – Whether the individual is married or in a committed partnership. Contacts with close friends and relatives – Frequency of social interactions with family and friends. Religious group membership – Whether the individual is part of a church or religious group and attends regularly. Participation in voluntary or community organizations – Involvement in social or civic groups outside of family and work. The SNI assigns points based on the individual's responses to the above components. Scores range from 0 to 4, with higher scores indicating greater social integration.
Non-standardised Questionnaire	A non-standardized questionnaire will be included to obtain information from the participants on their perceptions of the hearing aids. This includes Likert scale and open-ended questions.

3.5. Hypothesis and endpoint criteria

3.5.1. Primary endpoint criteria

The primary endpoint of this study is the self-reported outcome measured using the IOI-HA global score at 6 weeks post-fitting. Outcome measures will also be captured at 12-, 26-, and 52-weeks post-fitting. The endpoint assesses differences between study groups' overall hearing aid benefit and satisfaction.

Superiority

Null Hypothesis (H_0): There is no statistically significant difference in the IOI-HA global score between the experimental and control groups (statistically significant difference <3 points).

Alternative Hypothesis (H_1): The IOI-HA global score for one or both experimental groups is statistically significantly superior to the control group (statistically significant is represented by a margin of ≥3 points).

Non-inferiority

Null Hypothesis (H_0): Self-reported outcomes (IOI-HA) in the pre-set group are non-inferior to those in the smartphone-based in-situ fitting group, with the non-inferiority margin (δ_1) defined as 3.0 for the IOI-HA total score.

Alternative Hypothesis (H_1): Self-reported outcomes (IOI-HA) in the pre-set group are statistically significantly inferior to those in the smartphone-based in-situ fitting group, exceeding the predefined non-inferiority margin for the IOI-HA total score ($\delta_1 = 3.0$).

Clinical Relevance:

A difference of ≥3 points in the IOI-HA score is considered clinically significant based on prior research by Apple, representing a meaningful improvement in hearing aid benefit, satisfaction, and quality of life.

3.5.2. Exploratory variables

Exploratory variables are included in this study to gather additional insights and generate hypotheses for future research. Since these variables are primarily intended to uncover potential trends or relationships that are not yet well understood, formal hypotheses are specified. This allows for a more flexible analysis, where findings can be interpreted in a broader context without constrains of testing specific predictions.

4. Study population

4.1. Sample size determination and justification

This study will enroll a total of 90 participants (30 per study arm). Sample size calculations for this three-arm experimental study, involving four repeated measures (baseline and follow-up assessments at 6, 12, 26, and 52 weeks), were performed using the GLIMMPSE software (GNU General Public License, version 2). GLIMMPSE is specifically designed to compute sample size and statistical power for Generalized Estimating Equations (GEE), which are particularly suited for longitudinal and repeated measures data where observations within participants are correlated. In this study, GEE was deemed appropriate because it accounts for the within-subject correlations across multiple time points, ensuring more accurate and robust estimates for group effects. The calculations focused on the primary outcome, IOI-HA, and were designed to detect a medium effect size (d = 0.5; Cohen, 1969) with at least 80% statistical power. It is widely acknowledged that determining sample sizes to detect small effect sizes is often unnecessary, as such effects, while potentially statistically significant (p ≤0.05), may lack practical or real-world relevance (Baicus & amp; Caraiola, 2009; Peeters, 2016). Consequently, this study prioritized the detection of at least a medium effect size. The analysis indicated that a total of 69 participants (23 per arm) would be sufficient to achieve a statistical power of 0.811, assuming a Type I error rate of 0.05. The sample size calculation employed the Hotelling-Lawley Trace test and incorporated a correlation matrix with decreasing correlations over longer time intervals between measurements.

4.2. Inclusion and exclusion criteria

Inclusion criteria

Participants must meet all of the following criteria to be eligible for enrollment in the clinical trial:

- Aged 18 years or older
- Diagnosed with bilateral mild to severe hearing loss, defined as a pure-tone average (PTA) between 20 dB HL and <80 dB HL in both ears, based on WHO 2021 classification, confirmed during baseline assessments
- Willing and available to attend follow-up visits at 6 and 12 weeks post-fitting
- Able to provide written informed consent

Exclusion Criteria

Participants will be excluded from the study if they meet any of the following criteria:

- Younger than 18 years of age
- Hearing loss that is too severe (≥80 dB HL PTA) or normal hearing (<20 dB HL PTA)
- Unilateral hearing loss
- Presence of middle ear pathology, including otitis media or active ear drainage
- Unwilling or unavailable to commit to follow-up appointments at 6 and 12 weeks

4.3. Participant flow

Enrolment, inclusion, exclusion and progress for participants will be captured using the following type of CONSORT diagram.

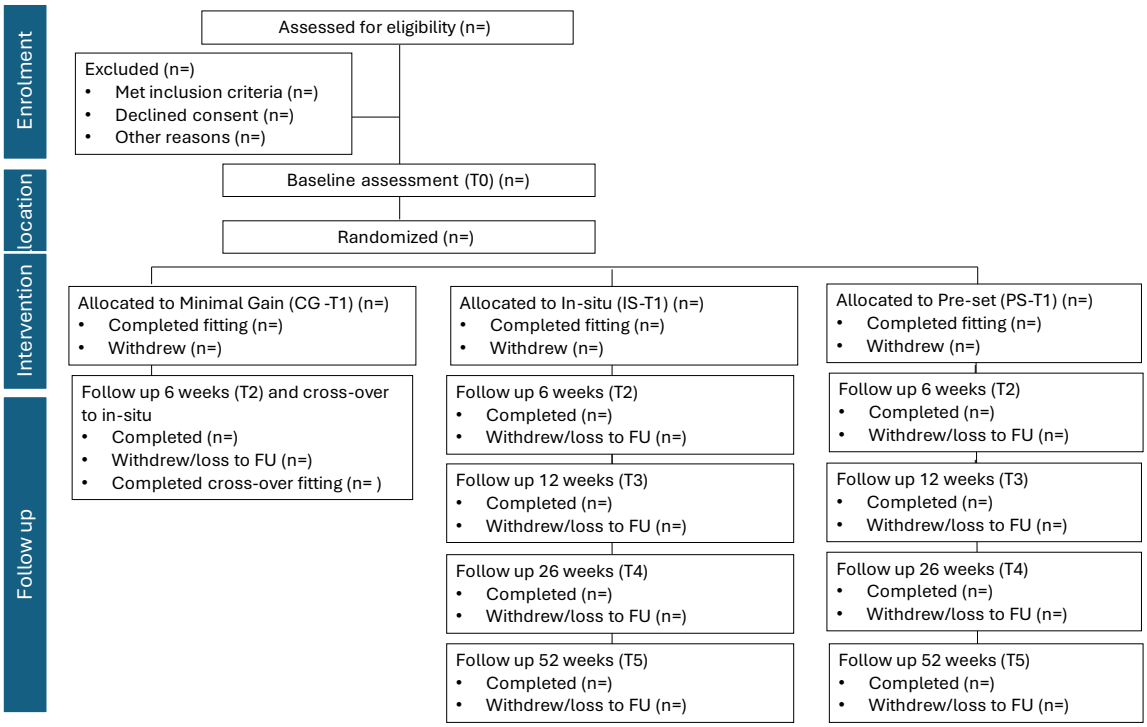

Figure 2. Participant flow diagram

Furthermore, Table 3 below outlines the timing of assessments for each variable included:

Table 3. Schedule of assessments

Assessment/ Procedure	**Baseline 6 (T0)**	**Fitting (T1)**	**6-week follow up (T2)**	**12-week follow up (T2)**	**26 week follow up (T3)**	**52 week follow up (T4)**
Informed consent	X					
Demographic information	X					
Otoscopy	X	X				
Pure tone audiometry	X					
Eligibility confirmation	X					
Randomization	X					
In-situ hearing test (Lumen Only)		X				
Hearing aid fitting		X				
RHHI-S		X	X	X	X	X
EQ-5D-5L			X	X	X	X

IOI-HA			X	X	X	X
SNI	X		X	X	X	X
Non-standardized questionnaire			X	X	X	X
Cross-over for control group			X			

5. Statistical analysis methods

5.1. Withdrawals, dropouts and handling of missing data

All randomized participants who receive at least one hearing aid fitting will be considered for inclusion in an Intention-to-Treat (ITT) analysis. The ITT principle ensures that participants are analyzed in the groups to which they were originally randomized, regardless of adherence, protocol deviation, crossover, or withdrawal. This approach preserves the benefits of randomization and provides an unbiased estimate of the real-world effectiveness of each intervention.

Participants who withdraw consent, are lost to follow-up, or are otherwise unable to complete the trial will remain in the ITT population, and their data will be included up to the point of withdrawal. Reasons for dropout will be documented, and descriptive analyses of attrition patterns will be provided by group and time point.

A comprehensive missing data analysis will be conducted after data cleaning to determine the extent and mechanism of missingness. Missing data will be classified into one of the following categories:

- Missing Completely at Random (MCAR)
- Missing at Random (MAR)
- Missing Not at Random (MNAR)
- Structurally Missing

The primary analysis will use a Linear Mixed-Effects Model (LMM), which is inherently robust to missing data under the MAR assumption. LMMs use maximum likelihood estimation to provide unbiased estimates without requiring imputation, provided that the probability of missingness depends only on observed data. This characteristic makes LMMs particularly well-suited to longitudinal clinical trials where some degree of missingness is expected.

Therefore, if missing data are determined to be limited and plausibly MAR, the primary analysis will proceed without multiple imputation, leveraging the LMM's robustness. However, if a substantial proportion of missing outcome data is identified, especially if missingness is differential across groups or correlated with baseline characteristics, then a formal ITT analysis incorporating multiple imputation will be conducted to account for potential bias and to enhance statistical power.

Multiple imputation will be carried out using appropriate predictive models, generating multiple datasets that are analysed separately and combined using Rubin's rules. The number of imputations will be guided by the proportion of missing data. Sensitivity analyses, including

complete case analysis and pattern mixture models, will be performed to test the robustness of findings under different assumptions about the missing data mechanism.

Exploratory analyses will be conducted to investigate additional factors that may influence hearing aid effectiveness and quality of life outcomes. Given the nature of these exploratory objectives, both within-group and between-group comparisons will be performed. As some analyses involve post-randomization factors or treatment crossover (e.g., within-group change in the control arm after crossover to in-situ fitting), the ITT principle will not be uniformly applied.

Instead, exploratory analyses will be conducted using an as-treated or per-protocol approach where appropriate, particularly for evaluating changes within individuals over time or the impact of adherence-related factors. Between-group exploratory comparisons that align with the original randomized assignments may still be analysed using the ITT population, where this preserves the validity of the comparison. All exploratory analyses will be considered hypothesis-generating and will be interpreted with appropriate caution. Results will not be used to draw confirmatory conclusions but will serve to inform future research and contextualize the primary and secondary findings.

5.2. Level of statistical significance

The overall significance level for hypothesis testing will be controlled at 5% ($\alpha = 0.05$). Although there are multiple primary comparisons, no corrections, such as the well-known Bonferroni correction, will be applied, as doing so has been criticised in the literature. The trouble with these types of corrections is that the significance changes with the number of comparisons. So, if two researchers are testing the same hypothesis, one could find a result to be significant, and the other not, depending on how many comparisons they make.

All secondary and exploratory analyses will be conducted using a 5% significance level, with results interpreted in an exploratory, hypothesis-generating context. No adjustments will be made for multiplicity in exploratory analyses for similar reasons mentioned above.

5.3. Statistical software

Analysis will be performed primarily using Statistical Packages of the Social Sciences (IBM SPSS v 30.0)

5.4. Participant characteristics and audiological variables

Description of the baseline characteristics will be presented by study arm. Discrete/factor variables will be summarized by frequencies and percentages. Percentages will be calculated according to the number of participants for whom the data are available. Continuous variables will be summarised by using mean and SD, and median and interquartile range (Q1-Q3).

- Age
- Sex
- Ability to use a phone
- State of readiness
- Self-reported years of hearing difficulty
- Otoscopy
- Pure tones average (based on audiogram performed by the audiologist)

- Ethnicity/ Race

5.5. Primary endpoint analysis

5.5.1. Testing the superiority hypothesis (Pre-set and In-situ versus control group)

The primary endpoint of this trial is the IOI-HA global score measured at 6 weeks post-fitting. Secondary observations of the IOI-HA score will be collected at 12, 26, and 52 weeks. The IOI-HA consists of 7 items, each scored on a 5-point ordinal scale, resulting in a global score ranging from 7 to 35. The IOI-HA will be treated as a continuous variable for modelling purposes, as is common in hearing aid outcome studies.

A Linear Mixed-Effects Model (LMM) will be used to evaluate the effect of treatment group over time on the IOI-HA global score. This modelling approach accounts for intra-individual correlation due to repeated measures. The LMM will include both fixed and random effects, specified as follows:

- Fixed effects:
  - Treatment group (categorical: In-situ, Pre-set, Control)
  - Time (categorical: 6, 12, 26, and 52 weeks)
  - Group × Time interaction, to test whether the effect of treatment differs over time

- Random effects:
  - Participant ID (random intercept), to account for repeated measurements within individuals

For the primary endpoint analysis, the model will be restricted to data collected at 6 weeks post-fitting, prior to crossover in the control group. This preserves the validity of the original randomization. Pairwise comparisons will be conducted between the in-situ group versus control and pre-set versus control.

Appropriate contrasts from the fitted linear mixed model will be used to calculate point estimates and two-sided 95% confidence intervals (CIs) for the difference in IOI-HA global scores between treatment groups. Superiority will be concluded if the lower bound of the 95% CI for the difference in means (Intervention - Control) exceeds the minimum clinically important difference (MCID) of 3 points on the IOI-HA global score. This threshold reflects the minimum score difference considered meaningful for patient benefit.

In addition to hypothesis testing, effect sizes will be calculated to quantify the magnitude of observed differences. For continuous outcomes such as the IOI-HA global score, Cohen's *d* will be used to estimate standardized mean differences between groups. Cohen's *d* will be calculated using the adjusted means and pooled standard deviations from the model. Values of *d* around 0.2, 0.5, and 0.8 will be interpreted as small, medium, and large effects, respectively. Reporting effect sizes alongside p-values and confidence intervals will aid in the interpretation of clinical relevance, particularly where statistical significance is marginal.

The formula for Cohen's *d* is as follows:

$$\mathrm{d} = \frac{M1 - M2}{SDpooled}$$

5.5.2. Testing the difference and non-inferiority between Pre-set and In-situ groups

To evaluate whether the Pre-set and In-situ hearing aid fittings result in statistically different self-reported outcomes, and to assess whether the Pre-set group is non-inferior to the In-situ group, a formal comparison between these two interventions will be conducted using the IOI-HA global score measured at 6 weeks post-fitting. This dual-purpose analysis allows for both the detection of meaningful differences and the demonstration that the lower-cost, pre-set solution performs comparably to the more customized in-situ fitting.

A Linear Mixed-Effects Model (LMM) will be used to assess the between-group difference. The model will be identical in structure to the primary endpoint analysis and will include:

- Fixed effects:
  - Treatment group (categorical: In-situ, Pre-set, Control)
  - Time (categorical: 6, 12, 26, and 52 weeks)
  - Group × Time interaction, to test whether the effect of treatment differs over time

- Random effects:
  - Participant ID (random intercept), to account for repeated measures within individuals

In addition, a non-inferiority analysis will be performed to determine whether the Pre-set group is not worse than the In-situ group by more than a clinically meaningful margin. The non-inferiority margin (Δ) is set at 3 points on the IOI-HA global score. Non-inferiority will be concluded if the lower bound of the 95% CI for the difference (Pre-set - In-situ) is greater than - 3.0 points. This threshold reflects the minimum difference considered clinically meaningful in self-reported hearing aid benefit.

5.6. Assessment of control group cross-over to in-situ

A within-subject analysis will be conducted to evaluate the change in outcomes among participants in the control group following crossover to the in-situ hearing aid fitting. Specifically, IOI-HA scores at 6 weeks (T2, minimal gain) and 12 weeks (T3, in-situ) will be compared. If data are normally distributed, a paired t-test will be used to assess the mean change in IOI-HA global score. If normality assumptions are violated, the non-parametric Wilcoxon signed-rank test will be applied. In addition to statistical significance, the mean change and corresponding 95% confidence interval will be reported, along with Cohen's *d* effect size to assess the magnitude of improvement. This analysis is exploratory and will provide insight into the benefit gained by the control group after receiving the active intervention.

5.7. Exploratory analysis of RHHI-S, EQ-5D-5L and SNI

The RHHI-S will be used as an exploratory outcome measure to assess participants' perceived emotional and social impact of hearing loss. The RHHI-S consists of 10 items, each scored

on a 3-point scale: 0 = “No”, 2 = “Sometimes”, and 4 = “Yes”, resulting in a total score ranging from 0 to 40. Higher scores reflect greater self-perceived hearing handicap. Standard interpretive ranges will be used for descriptive purposes:

- 0-8 = No handicap
- 10-24 = Mild to moderate handicap
- 26-40 = Significant handicap

Exploratory analyses will be conducted to assess differences in RHHI-S scores between the three study arms (In-situ, Pre-set, Control) at each follow-up time point (6, 12, 26, and 52 weeks). The RHHI-S will be treated as a continuous variable for modeling purposes.

A Linear Mixed-Effects Model (LMM) will be used to evaluate group-level differences in RHHI-S scores over time while accounting for repeated measures within participants. The model will include:

- Fixed effects:
    - Treatment group (In-situ, Pre-set, Control)
    - Time (6, 12, 26, and 52 weeks)
    - Group × Time interaction

- Random effects:
    - Participant ID (random intercept), to account for within-subject correlation

Pairwise comparisons of estimated marginal means will be conducted between groups at each time point. A two-sided alpha level of 0.05 will be used for exploratory comparisons. Adjustments for multiple comparisons (e.g., Tukey's method) may be applied as appropriate.

Descriptive statistics will be used to summarize the EQ-5D-5L data across the three study groups (In-situ, Pre-set, and Control) at each follow-up time point (6, 12, 26, and 52 weeks). The EQ-5D-5L instrument comprises two components: a descriptive system and a visual analogue scale (VAS). The descriptive system captures participant ratings across five health domains, mobility, self-care, usual activities, pain/discomfort, and anxiety/depression, each scored on a five-level scale ranging from "no problems" to "extreme problems." These responses will be converted into a single utility index score using the validated Ugandan value set**,** given the absence of a South African-specific tariff. This approach has been successfully applied in related South African health research and ensures culturally and regionally appropriate interpretation of health-related quality of life. The VAS component, which reflects a participant’s self-rated overall health on a scale from 0 (worst imaginable health) to 100 (best imaginable health), will be summarized using means, medians, standard deviations, and interquartile ranges by group and time point. In addition, a cut-off score of ≥73 will be applied to the VAS, as per Moyo et al. (2023), to categorize participants into those with a “good” versus “poor” perceived health state. The proportion of participants exceeding this threshold will be reported descriptively for each group and follow-up time point. These descriptive analyses will complement the inferential models by providing a clinically meaningful understanding of both self-rated and utility-based health outcomes over time, supporting the interpretation of potential improvements associated with the hearing aid interventions.

To assess between-group differences in EQ-5D-5L scores, Linear Mixed-Effects Models (LMMs) will be employed, treating treatment group**,** time**,** and the group × time interaction as fixed effects, and participant ID as a random intercept to account for repeated measures. This model allows for evaluation of both group differences at specific time points and changes over time while accounting for within-subject correlation. Additionally, at each time point, one-way ANOVA will be used to compare the EQ-5D-5L utility index and VAS scores across the three groups if data meet parametric assumptions. In cases where normality or homogeneity of variance is not met, the Kruskal-Wallis test will be used as a non-parametric alternative. For significant results, post hoc pairwise comparisons will be performed using Tukey's HSD or Dunn's test, as appropriate, with adjustments for multiple comparisons.

The Berman-Syme Social Network Index (SNI) evaluates participants' social connectedness based on four domains: (1) marital status, (2) frequency and quantity of contact with friends and relatives, (3) participation in religious meetings, and (4) involvement in community or organizational groups. Following the approach outlined by Loucks et al. (2006), responses on the Berkman-Syme SNI are converted into binary scores and then summed to create a composite index ranging from 0 to 4. Participants receive a score of 1 if they are married and 0 if they are not. For close social contacts, a score of 0 is assigned if an individual reports having 0–2 close friends and 0–2 close relatives; otherwise, a score of 1 is given. Regarding participation in community organizations, individuals who do not participate receive a score of 0, while those who do are assigned a score of 1. For religious attendance, a score of 0 is given to those attending services less than or equal to every few months, and a score of 1 is assigned to those attending once or twice a month or more frequently. The total score reflects the degree of social integration, with higher scores indicating greater social connectedness. The scores will be related to the demographic questions (age, gender, degree of hearing loss) using correlations and tests for differences. As explained above, the tests for differences between two unrelated/independent groups will be conducted using the independent samples t-test (if normal) or the Mann-Whitney test (if non-normal). For three or more groups, the one-way ANOVA test (if parametric) or the Kruskal-Wallis test (if non-normal) will be used. Predictors of higher or lower social connectedness scores will be explored using generalized linear models, with age, gender and degree of hearing loss included as covariates.

6. Data Management

6.1. Overview of data collection methods

Data for this trial will be collected prospectively at multiple time points to ensure completeness, consistency, and integrity. Assessments will take place at baseline (T0), hearing aid fitting (T1), and subsequent follow-up visits (T2-T5). At each visit, participants will complete standardized self-report questionnaires including the IOI-HA, RHHI-S**,** and EQ-5D-5L. Questionnaires will be administered either electronically or on paper by trained community healthcare workers (CHWs) or administrative staff to ensure standardization and minimize missing or inconsistent responses.

6.2. Data entry, cleaning and validation process

Manual data entry from paper-based questionnaires and audiometric records will be performed by trained research personnel using a secure data entry platform hosted on the University of Pretoria’s UP Drive. To promote data accuracy, a random audit procedure will

be implemented whereby approximately 30% of manually entered data will be rechecked by a second reviewer not actively involved with the data collection process. Discrepancies identified during auditing will be resolved through reference to original source documents, and all changes will be logged in an audit trail.

Following data entry, the dataset will undergo a structured data cleaning process, which includes checks for missing values, implausible entries (e.g., out-of-range scores or illogical date sequences), and duplicates. Logical validation rules will be applied to ensure that temporal relationships between study visits and responses are consistent. Any corrections to the dataset will be recorded with metadata specifying the change, rationale, and responsible team member.

Once cleaned, data validation checks will be performed to ensure readiness for statistical analysis. This will include confirming that all expected cases are present, that primary and secondary outcome variables are fully populated, and that the format and structure of the dataset align with the requirements of the statistical analysis plan. Initial data summaries (e.g., distributions and descriptive statistics) will be used to inspect variable quality and detect any anomalies requiring further review.

6.3. Software and tools used for data analysis

All statistical analyses will be conducted using IBM SPSS Statistics (v30.0). This software will be used for generating descriptive summaries, performing inferential statistics, modeling repeated measures data via linear mixed-effects models, and handling both parametric and non-parametric tests as outlined in the statistical analysis plan. In cases where additional flexibility or visualization is required (e.g., for exploratory or graphical outputs), supplementary use of R (version 4.3.2) may be considered.

All data analyses will be carried out on password-protected computers located within the secured study network. Data files will be backed up regularly and stored on encrypted institutional servers with access restricted to authorized study personnel only. Final datasets will be locked prior to analysis, and all intermediate versions will be archived with date stamps to ensure data traceability and reproducibility.

7. Reporting of results

7.1. Format for reporting primary and exploratory outcomes

Results will be reported using a clear and standardized structure to facilitate interpretation, assess statistical robustness, and highlight clinically meaningful differences between the Pre-set, In-situ, and Control (Minimal Gain) groups. For all primary, secondary, and exploratory outcomes, the following elements will be reported:

- Point Estimates and Mean Differences**:** Primary outcomes (e.g., IOI-HA global scores at 6 weeks) and secondary outcomes (e.g., RHHI-S, EQ-5D-5L utility index and VAS scores) will be summarized by treatment group using estimated marginal means derived from linear mixed-effects models. Mean differences between groups will be reported, including those used to assess superiority and non-inferiority hypotheses.

- Confidence Intervals (CIs): All estimates of between-group differences will be accompanied by 95% confidence intervals, providing a measure of precision and allowing for interpretation relative to predefined clinical margins (e.g., the 3-point superiority and non-inferiority margin on the IOI-HA scale).
- P-values**:** Two-sided p-values will be reported for hypothesis testing. For the non-inferiority comparison between Pre-set and In-situ groups, p-values will be supplemented by the interpretation of the confidence interval relative to the non-inferiority margin. For exploratory analyses and secondary outcomes, p-values will be interpreted with caution and contextualized alongside effect sizes and confidence intervals.
- Effect Sizes: Cohen's *d* will be reported for key continuous outcomes to convey the standardized magnitude of between-group differences. Thresholds for interpretation will follow conventional guidelines: 0.2 (small), 0.5 (medium), and 0.8 (large).

7.2. Graphical presentation of results

To support interpretation and improve communication of the results, findings will also be presented visually using appropriate graphical formats (examples provided below):

- Bar and line graphs will be used to illustrate the change in mean outcome scores (e.g., IOI-HA, RHHI-S, EQ-5D-5L) across time points for each intervention group. Error bars representing 95% confidence intervals will be included to show variability and precision.
- Box plots will be used to display the distribution of continuous variables such as IOI-HA, VAS scores, and data logging metrics (e.g., average daily usage), providing insights into medians, interquartile ranges, and potential outliers by group and time point.
- Scatterplots with fitted regression lines may be used to explore associations between continuous outcomes (e.g., EQ-5D-5L VAS vs. PTA or hearing aid usage), especially in exploratory analyses. For non-inferiority comparisons, confidence interval plots against the predefined margin (-3 points) may be included to visually demonstrate the decision boundary.

7.3. Interpretation of findings in relation to study objectives

Interpretation of findings will be guided by the study's primary objectives of (1) determining whether the Pre-set and In-situ groups are superior to the Control group in terms of IOI-HA scores, and (2) whether the Pre-set group is non-inferior to the In-situ group. Superiority will be declared if the lower bound of the 95% confidence interval for the mean difference between an intervention group and control exceeds 3 points. Non-inferiority will be concluded if the lower bound of the CI for the Pre-set - In-situ comparison exceeds -3 points.

Beyond statistical significance, the interpretation will emphasize clinical relevance by assessing whether observed differences meet or exceed minimum clinically important differences (MCIDs) for each measure. The magnitude of effect sizes, consistency of trends across time points, and the directionality of change will be considered when interpreting both confirmatory and exploratory outcomes.

Any carry-over effects, deviations from model assumptions, or missing data patterns identified during analysis will be transparently reported and their potential impact on findings discussed. In particular, any influence of attrition or non-adherence will be examined through sensitivity analyses. The results of exploratory analyses (e.g., on self-efficacy, digital literacy, and hearing aid usage patterns) will be contextualized as hypothesis-generating, and implications for clinical practice and future research will be highlighted where appropriate.

8. Reporting of adverse events

8.1. Identification and classification of adverse events

All adverse events (AEs) reported by participants during the trial period will be systematically documented, reviewed and analyzed to ensure participant safety and to monitor the risk profile of each intervention. Adverse events will be classified into the following two categories:

- *Serious adverse events (SAEs):* Defined as any untoward medical occurrence that results in death, is life-threatening, requires inpatient hospitalization or prolongation of existing hospitalization, results in persistent or significant disability/incapacity, or any other condition deemed serious by the principal investigator.
- *Non-Serious Adverse Events:* Any medical occurrence that does not meet the criteria for an SAE but may be temporally associated with the study intervention (e.g., minor skin irritation, discomfort from hearing aid use, or difficulty adjusting to amplification).

Each adverse event will be assessed for:

- Severity (mild, moderate, or severe)
- Causality (definitely related, probably related, possibly related, unlikely related, or unrelated to the study intervention)

Standard definitions and classification criteria will be used in alignment with Good Clinical Practice (GCP) and institutional SOPs.

8.2. Data Collection and Documentation

Adverse events will be recorded in a dedicated section of the participant's Case Report Form (CRF)**.** For each event, the following details will be documented:

- *Event Description:* A narrative summary of the event, including onset date, resolution date, duration, and any associated symptoms.
- *Severity Rating:* Categorized as mild (transient, no treatment needed), moderate (interferes with activities, may require minimal intervention), or severe (significant disruption of functioning or requiring medical attention).
- *Outcome:* Whether the participant recovered, is still experiencing the event, or had long-term consequences.
- *Relationship to Study Intervention:* Determined by the principal investigator using clinical judgment.
- *Action Taken:* Includes changes to study procedures, discontinuation, or medical referral.

All adverse event records will be retained in the study's source documentation and trial master file and included in periodic safety reports.

8.3. Reporting procedures

Adverse events will be reported in accordance with institutional guidelines and regulatory requirements:

- *Serious Adverse Events (SAEs):* All SAEs will be reported to the Humanities Research Ethics Committee at the University of Pretoria a within 24 hours of the site becoming aware of the event. A written follow-up report containing updated clinical information and resolution status will be submitted within 7 calendar days.
- *Non-Serious Adverse Events***:** Non-serious AEs will be documented in real time and reported in aggregate summaries at regular intervals to the DMC and included in progress reports to the HREC. The summaries will include the number, type, severity, and presumed relationship to study interventions for each treatment group.

Any adverse event deemed unexpected or related to device malfunction (e.g., ear canal injury, persistent discomfort) will be reviewed by the trial safety officer and, if necessary, reported to the device manufacturer under post-market surveillance obligations.

8.4. Data Analysis and Interpretation

Adverse event data will be descriptively analyzed to assess the safety profile of the three study interventions: In-situ, Pre-set, and Control (Minimal Gain). The frequency and types of AEs will be summarized as counts and proportions, stratified by treatment group and time point. Events will also be grouped by system organ class and severity.

Comparative analysis will be conducted to explore whether the incidence of AEs differs meaningfully between groups. Particular attention will be given to:

- Reports of discomfort, irritation, or device-related issues
- Serious events requiring medical follow-up
- Any trends associated with specific device models or fitting methods

If warranted, inferential statistics (e.g., Chi-square tests) may be used to test for significant group differences in adverse event occurrence. The relationship between adverse events and study interventions will be reviewed to identify any potential causal links**,** and these will be discussed in the final trial report alongside recommendations for hearing aid safety monitoring in similar low-resource settings.

9. References

Baicus, C., & Caraiola, S. (2009). Effect measure for quantitative endpoints: Statistical versus clinical significance, or "how large the scale is?". *European Journal of Internal Medicine, 20*(5), e124–e125. https://doi.org/10.1016/j.ejim.2008.10.002

Bisgaard, N., Zimmer, S., Laureyns, M., & Groth, J. (2021). A model for estimating hearing aid coverage world-wide using historical data on hearing aid sales. *International Journal of Audiology, 60*(10), 761–769. https://doi.org/10.1080/14992027.2021.1962551

Cohen, J. (1969). *Statistical power analysis for the behavioral sciences*. Academic Press.

Cox, R. M., & Alexander, G. C. (2002). The International Outcome Inventory for Hearing Aids (IOI-HA): Psychometric properties of the English version. *International Journal of Audiology, 41*(1), 30–35. https://doi.org/10.3109/14992020209101309

Herdman, M., Gudex, C., Lloyd, A., Janssen, M. F., Kind, P., Parkin, D., Bonsel, G., & Badia, X. (2011). Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Quality of Life Research, 20*(10), 1727–1736. https://doi.org/10.1007/s11136-011-9903-x

Keidser, G., Dillon, H., Flax, M., Ching, T., & Brewer, S. (2011). The NAL-NL2 prescription procedure. *Audiology Research, 1*(1), 88–90. https://doi.org/10.4081/audiores.2011.e24

Loucks, E. B., Sullivan, L. M., D'Agostino, R. B., Sr., Larson, M. G., Berkman, L. F., & Benjamin, E. J. (2006). Social networks and inflammatory markers in the Framingham Heart Study. *Journal of Biosocial Science, 38*(6), 835–842. https://doi.org/10.1017/s0021932005001203

Mulwafu, W., Ensink, R., Kuper, H., & Fagan, J. (2017). Survey of ENT services in Sub-Saharan Africa: Little progress between 2009 and 2015. *Global Health Action, 10*(1), 1289736. https://doi.org/10.1080/16549716.2017.1289736

Peeters, M. J. (2016). Practical significance: Moving beyond statistical significance. *Currents in Pharmacy Teaching and Learning, 8*(1), 83–89. https://doi.org/10.1016/j.cptl.2015.09.001

World Health Organization. (2021). *World report on hearing*. https://www.who.int/publications/i/item/world-report-on-hearing

World Health Organization. (2024). *Hearing aid service delivery approaches for low-and middle-income settings*. World Health Organization. https://www.who.int/publications/i/item/9789240092936