Plan Overview

A Data Management Plan created using DMPonline

Title: Developing norm-values for High-frequency brainstem Evoked Response Audiometry in young children

Creator: Marieke Snijder-van As

Principal Investigator: R.A. Boerboom

Data Manager: Dax Steins

Project Administrator: K.S. Rhebergen

Affiliation: UMC Utrecht

Template: UMC Utrecht DMP

Project abstract:

For early identification of ototoxic hearing loss in young, non-cooperating children, we aim to develop an objective method to determine high-frequency (HF) hearing loss using the auditory brainstem responses (ABR). This includes the selection and optimization of several HF stimuli for ABR, correlation with default tone audiometry, and determination of norm response values after the pilot phase.

ID: 77996

Start date: 01-06-2021

End date: 01-06-2023

Last modified: 05-07-2021

Copyright information:

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customise it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

Developing norm-values for High-frequency brainstem Evoked Response Audiometry in young children

1. General features

1.1. Please fill in the table below. When not applicable (yet), please fill in N/A.

DMP template version	29 (don't change)
ABR number *(only for human-related research)*	
METC number *(only for human-related research)*	TBC
DEC number *(only for animal-related research)*	
Acronym/short study title	HERA
Name Research Folder	xx-xxx_HERA
Name Division	Surgical Specialties
Name Department	ENT department
Partner Organization	
Start date study	1-6-2021
Planned end date study	1-6-2023
Name of datamanager consulted*	Dax Steins
Check date by datamanager	14-06-2021

1.2 Select the specifics that are applicable for your research.

- Non-WMO
- Use of Questionnaires
- Monocenter study
- Clinical study
- Interventional study

2. Data Collection

2.1 Give a short description of the research data.

The project is described in (20210608) UMCU_Onderzoeksprotocol nWMO (v0.7). Parts between brackets might be updated.

Project aim: selection and optimization of several HF stimuli for ABR, correlation with default tone audiometry, and determination of norm response values after the pilot phase.

Primary Objective: Selection and optimisation of HF stimuli (both for head phones and bone conduction transducer) for ABR measurements, in order to obtain a stimulus that gives the best signal response (regarding peak detection and SNR) and is best correlated with behavioural tone audiometry

Secondary Objective(s): latency of peak V.

Study population: healthy volunteers (18-25 years) with normal hearing.

Study procedure:

1. Filling out questionnaire (F1. Vragenlijst voor proefpersonen 20210528) to obtain participants' (hearing specific) characteristics - on paper => scanned to PDF

2. Hearing screening according to the Dutch health care protocol. The following frequencies will be tested: 0.25, 0.5, 1, 2, 4, 6, 8, 10, 12.5, 14, 16 kHz - done in the default clinical UMCU setting, probably with Audiqueen (we are currently in a transition from using DECOS to using Audiqueen). Output: print (on papier), from which thresholds at the different frequencies in dBHL/dBSPL can be substracted.

3. Impedance measurement with a calibrated tympanometer at the UMCU, also in Audiqueen. Output: print (on paper)

4. ABR tests with different types of stimuli and at different levels. Stored in database of SmartEP software of the Duet ABR clinical device. The peaks in the detected signals are marked. Output: PDF and ASCI-file with both raw-signals and marked peak positions (amplitude, latency).

All data is saved in the various datasystems with the anonymous internal study number (HERApat_XXX). After datacollection, data will be stored in PDF of asci in de research datafolder.

Quantitative data is (finally) stored in Excel tables. Analysis will be done in Excel, SPSS and/or Matlab.

Final data:

Subjects	Volume	Data Source	Data Capture Tool	File Type	Format	Storage space
Human	40	Experimental (extracted data of peaks)	Excel	Quantitative	csv	0-10 GB
Human	40	Experimental (exported PDFs with measured curves)		Tekst/Image	pdf	0-10 GB
Human	40	Questionnaire		Tekst/image	pdf	0-10 GB
Human	40	Experimental (hearing thresholds toonaudio)		Tekst/image	pdf	0-10 GB
Human	40	Experimental (extracted toonaudio data)	Excel	Quantitative	csv	0-10 GB
Human	40	Experimental (tympanometrie)		Image	pdf	0-10 GB

2.2 Do you reuse existing data?

- No, please specify

Literature review showed that, although some research data exists about the technological aspects of the subject, there is little data available to answer our specific research question. There is thus a need to collect primary data on this topic. Because these HF-ABR-signals have not been used @UMCU before, there is no relevant data available in HIX.

2.3 Describe who will have access to which data during your study.

Upon informed consent by the subject, a research ID will be provided. The research ID is noted on the informed consent and stored seperately. All data collection will be done with the research ID.

Type of data	Who has access
Direct identifying personal data	Research team, Datamanager
Pseudonymized data	Research team, datamanager
Key table linking study specific IDs to Patient IDs	Research team, Datamanager

2.4 Describe how you will take care of good data quality.

All data is stored as PDF (questionnaire, tone audiometry and ABR measurement with indicated peaks), this can be seen as a backup/frozen dataset. All quantitative data (thresholds and peak-though values) will we saved in an csv-file. Data will be matched by study subject code.

A CRF on paper is made with validation checks to ensure data collection is according to predefined standards.

Peak/though marking is checked by another member of the research team, to ensure the data quality.

#	Question	Yes	No	N/A
1.	Do you use a certified Data Capture Tool or Electronic Lab Notebook?		X	
2.	Have you built in skips and validation checks?	X		
3.	Do you perform repeated measurements?		X	
4.	Are your devices calibrated?	X		
5.	Are your data (partially) checked by others (4 eyes principle)?	X		
6.	Are your data fully up to date?			X
7.	Do you lock your raw data (frozen dataset)	X		
8.	Do you keep a logging (audit trail) of all changes?		X	
9.	Do you have a policy for handling missing data?		X	
10.	Do you have a policy for handling outliers?		X	

2.5 Specify data management costs and how you plan to cover these costs.

#	Type of costs	Division ("overhead")	Funder	Other (specify)
1.	Time of datamanager	X		
2.	Storage	X		
3.	Data capture tool (audiometry and ABR)	X		
4.	Medical devices	X		
5.	Design of paper CRF	X		

2.6 State how ownership of the data and intellectual property rights (IPR) to the data will be managed, and which agreements will be or are made.

UMC Utrecht is the owner of all collected data for this study. Analysed data will be published. Selected and optimized HF stimuli will be made available for use in other centra.

3. Personal data (Data Protection Impact Assessment (DPIA) light)

Will you be using personal data (direct or indirect identifying) from the Electronic Patient Dossier (EPD), DNA, body material, images or any other form of personal data?

- Yes, go to next question

I will process personal data. I have consulted the division datamanager and I do not have to complete a full DPIA. I therefore fill out this DPIA light and proceed to 3.1.

3.1 Describe which personal data you are collecting and why you need them.

Which personal data?	Why?
Demographic data (age, gender)	To describe study population
Medical history	For verifiying inclusion criteria
Name and email address of participants	To be able to communicate on participation. (NB: stored seperately from other data, data is only processed pseudonymized)
Audiological data (thresholds and ABR peak detection)	Study objective: determine correlations and norm values.

3.2 What legal right do you have to process personal data?

- Study-specific informed consent

3.3 Describe how you manage your data to comply to the rights of study participants.

The data are pseudonymized and the linking table to personal data is saved. An authorized person manages the linking table, can re-identify study participants when necessary and deliver, correct or delete the data. The procedure can be found in the research folder of the division.

Right	
Right of Access	Research data are coded, but can be linked back to personal data, so we can generate a personal record at the moment the person requires that. This needs to be done by an authorized person.
Right of Objection	We use informed consents.
Right to be Forgotten	In the informed consent we state that the study participant can stop taking part in the research. Removal of collected data from the research database will only be granted as long as it is still possible to enroll new participants in the study. After that the removal of the collected data for the research database cannot be granted because this would result in a research bias.
Right of Rectification	The authorized person will give the code for which data have to be rectified.

3.4 Describe the tools and procedures that you use to ensure that only authorized persons have access to personal data.

We use the secured Research Folder Structure that ensures that only authorized personnel has access to personal data, including the key table that links personal data to the pseudoID.

3.5 Describe how you ensure secure transport of personal data and what contracts are in place for doing that.

We will not transport any personal data outside the UMCU network drives.

4. Data Storage and Backup

4.1 Describe where you will store your data and documentation during the research.

The digital files will be stored in the secured Research Folder Structure of the UMC Utrecht. We will need <50 GB storage space, so the capacity of the network drive will be sufficient. Paper dossiers will be stored safely in a locked cabinet in a locked room in the UMC Utrecht. A project specific procedure is in place for access to the paper dossiers. Documentation of this procedure is stored in the Research Folder Structure.

4.2 Describe your backup strategy or the automated backup strategy of your storage locations.

All (research) data is stored on UMC Utrecht networked drives from which backups are made automatically twice a day by the division IT (dIT).

5. Metadata and Documentation

5.1 Describe the metadata that you will collect and which standards you use.

We do not use metadata standards yet.
We use a CRF on paper to standardize the data collection and log how the procedure for the specific participent went (i.e. researcher, order of stimuli, any deviations from the test protocol).
In read-me files and headings of the columns in the excel-file it is noted what data is presented.

5.2 Describe your version control and file naming standards.

We will distinguish versions by indicating the version in the filename by adding a code after each edit, for example V1.1 (first number

for major versions, last for minor versions) and/or a saving date in the filename. The most recent copy is always used as the source, and before any editing, this file is saved with the new version/date code in the filename. The file with the highest code number is the most recent version. Upon saving a newer version, we will move previous versions to a folder OLD.
Filenames will be choosen in accordance with the research data structure of the UMCU.

6. Data Analysis

6 Describe how you will make the data analysis procedure insightful for peers.

Primary data are the recorded potential signals. In the manufacturer software peaks and troughs will be marked by two people (independently or one person checking the marking of the other). The corresponding peak amplitude and SNR values will be exported and are a secondary data-set which can still be seen as 'RAW-data'.
I will make an overview of these secondary datasets and the analysis scripts will contain comments, such that it is fully clear how the statistical analysis is performed. Peers will be able to repeat the analysis based on my overview.

7. Data Preservation and Archiving

7.1 Describe which data and documents are needed to reproduce your findings.

The data package will contain: the raw data, the study protocol describing the methods and materials, the script to process the data, the scripts leading to tables and figures in the publication, a codebook with explanations on the variable names, and a ‘read_me.txt’ file with an overview of files included and their content and use.
After finishing the project, this documentation will be stored at the UMC Utrecht research drive and is under the responsibility of the Principal Investigator of the research group

7.2 Describe for how long the data and documents needed for reproducibility will be available.

Data and documentation needed to reproduce findings from this non-WMO study will be stored for at least 15 years

7.3 Describe which archive or repository (include the link!) you will use for long-term archiving of your data and whether the repository is certified.

After finishing the project, the data package will be stored at the UMC Utrecht Research Folder Structure and is under the responsibility of the Principal Investigator of the research group. When the UMC Utrecht repository is available, the data package will be published here.

7.4 Give the Persistent Identifier (PID) that you will use as a permanent link to your published dataset.

I will be using a DOI-code and will update this plan as soon as I have the code.

8. Data Sharing Statement

8.1 Describe what reuse of your research data you intend or foresee, and what audience will be interested in your data.

Our research data is most likely not of interest to others, the used/selected stimuli and the developed normvalues will be. These will be available upon publication.

8.2 Are there any reasons to make part of the data NOT publicly available or to restrict access to the data once made publicly available?

- No, all data generated in this project will be made publicly available without any restrictions

Our data will be shared with third parties upon request and after approval of the Principle Investigator. The criteria and time period will be determined on a case-by-case basis.

8.3 Describe which metadata will be available with the data and what methods or software tools are needed to reuse the data.

Along with the publication, the codebook of the data and scripts of analysis in SPSS/Matlab/R/Python will be available upon request.

8.4 Describe when and for how long the (meta)data will be available for reuse

- (Meta)data will be available as soon as article is published

8.5 Describe where you will make your data findable and available to others.

The data is stored on the research disk and upon request to the principel investigator it can be shared.