**Title**: Detecting and Disambiguating “Hidden” Cochlear Pathologies

**Background:** Some types of damage to the cochlea are not detected by standard, clinical hearing tests. In this study, I investigate two such “hidden” pathologies, inner hair cell (IHC) damage and cochlear synaptopathy (CS), which can masquerade as clinically normal hearing despite potentially having detrimental effects on suprathreshold listening and speech perception. This dataset aims to identify whether other biomarkers of auditory function are more sensitive to these two pathologies than the traditional hearing test, and if any of these biomarkers differentiate between IHC damage and CS.

**Data Description:** 16 chinchillas were randomly assigned to one of two exposure groups, maintaining an equal number of male and female chinchillas in each group. One group was exposed to a 1000 Hz octave-band noise for 2 hours to induce cochlear synaptopathy. This group is labeled TTS because the noise exposure causes a temporary threshold shift (TTS) that recovers after two weeks. The other group was administered 38 mg/kg i.p. Carboplatin (CA) to induce mild (~15%) selective inner-hair-cell loss and significant stereocilia dysfunction in the surviving inner hair cells.

Both before and 2 weeks after the exposure, hearing was evaluated using 6 different biomarkers:

1. Auditory Brainstem Response (ABR) Thresholds: tone bursts at different frequencies elicit an electrical response which is measured from subdermal needle electrodes. The lowest sound level which elicits a repeatable response is deemed to be the hearing threshold. Thresholds were measured at 500, 1000, 2000, 4000, and 8000 Hz. This is the gold standard for clinically assessment of hearing loss.
2. Envelope Following Response (EFR): the electrophysiological response to a rectangular amplitude modulated (RAM) tone was recorded. The EFR amplitude was calculated as the sum of the first four harmonics of the modulation frequency.
3. Middle Ear Muscle Reflex (MEMR) Threshold: The lowest noise level that elicited a contraction of the middle ear muscle was recorded.
4. Distortion Product Otoacoustic Emissions (DPOAEs): DPOAEs were measured from 500-16000 Hz and reported as the weighted average at 9 discrete half-octave frequency bands from 750-12000 Hz.

All data was collected at Purdue University as part of my thesis project.

**Aims of the experiments**: The primary goal of this study is to investigate biomarkers for two different inner ear pathologies typically hidden from standard hearing testing. I will use this dataset to answer the following questions:

1. Are any of the biomarkers (ABR, EFR, MEMR, DPOAE) sensitive to inner hair cell damage or cochlear synaptopathy? i.e., are pre-exposure responses different from post-exposure responses for these measures?
2. Does IHC damage affect the biomarkers differently than CS? i.e., are post-exposure responses for the two groups different across any of the individual biomarkers?
3. Can integration of results across biomarkers better differentiate the groups than a single biomarker alone?

**Hypotheses:** Following the above aims, I hypothesize that:

1. ABR and DPOAEs will not be different before and after exposure since they are primarily driven by outer hair cell function rather than inner hair cell function. EFR amplitudes and MEMR thresholds will both be significantly worse in the post-exposure condition compared to the pre-exposure condition.
2. While both CS and IHC dysfunction will both reduce EFR amplitudes, MEMR thresholds will be reduced more in the CS group than the IHC group.
3. Integration across all biomarkers will differentiate between the IHC dysfunction and CS groups better than any individual metric.

**Proposed methods**: I have the full dataset organized into a single CSV that will need to be modified to appropriately analyze and plot the data (see example code for summary figures). To address the goals of this project, I will need to compare the outcomes of a few different measures. What R packages/functions do you plan to use? Packages I plan to use are: Tidyverse, ggplot2, and XXX. What kind of figures do you plan to communicate the findings? What statistical methods do you plan to test the hypotheses?

**Simple summary stats of the data:** I have included a few plots below to give a sense of the data and the types of analyses that will be completed

A graph of different types of lines

Description automatically generated with medium confidenceFIGURE 1: Average hearing thresholds for the two groups before (black) and after (red) exposure. One animal showed a significant elevation in thresholds after the CA exposure and one animal in the TTS group has elevated thresholds at baseline.

**A graph of different types of lines

Description automatically generated with medium confidence**

FIGURE 2: Average DPOAE amplitudes for the two groups before (black) and after (red) exposure. Overall, there is not much difference between pre vs post exposure. If anything, there may be a slight increase in amplitudes in the Carboplatin (CA) group and a slight decrease in amplitudes in the (TTS)

**A graph with red and black squares

Description automatically generated**FIGURE 3: Average EFR amplitudes for the two groups before (black) and after (red) exposure.A graph of different types of lines

Description automatically generated with medium confidence In general, both CA and TTS groups seems to show a reduction in EFR amplitudes after exposure, though it may be worse in the CA group than the TTS group.

**GitHub Repository:** All code for this project (including the demo figures above) can be found at this link:

**References**

Include relevant literature for the background of your research question.

Include relevant literature that generates or utilizes the datasets

[format: single space, 1inch margin, text+figures/tables no more than 2 pages, references can go over the two-page limit]

-brief description of the experiments

-what data is collected (measurements? Pictures? Sound/video Recordings? Sequences etc.)

-What taxa are involved?

-Location

-Duration, intervals of the experiments

-Number of relational tables (analyze at least two tables), dimensions of the data

List questions the authors/you plan to answer.