Applications of Machine Learning in Retrospective Studies on Hearing

François Charih, Ashlynn Steeves, Matthew Bromwich, Renée Lefrançois
Amy E. Mark and James R. Green

October 29th, 2018
IEEE Life Sciences Conference 2018

CU BIC CARLETON UNIVERSITY BIOMEDICAL INFORMATICS COLABORATORY
Carleton UNIVERSITY Canada's Capital University
clearwater

Background

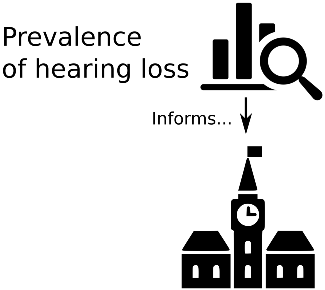
Background

Studies on the prevalence of hearing loss are important

Prevalence
of hearing loss

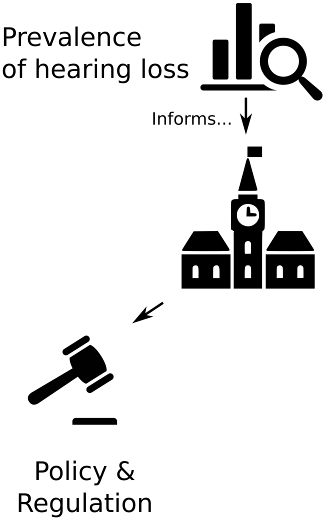
Background

Studies on the prevalence of hearing loss are important



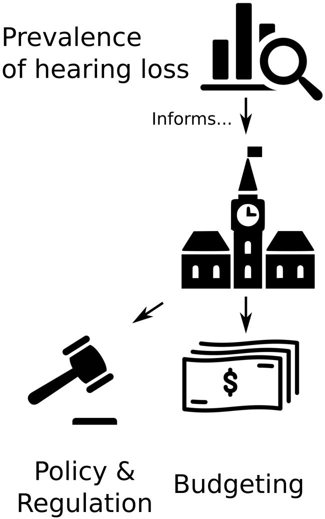
Background

Studies on the prevalence of hearing loss are important



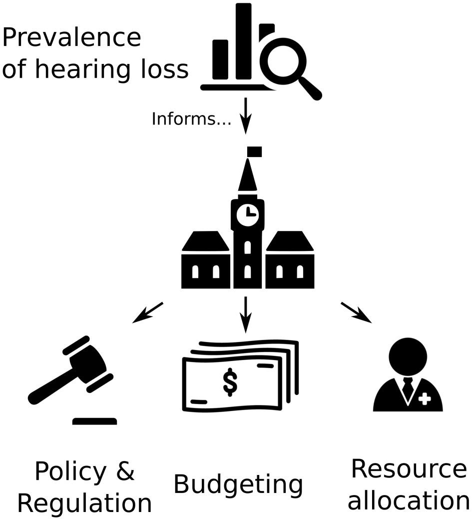
Background

Studies on the prevalence of hearing loss are important



Background

Studies on the prevalence of hearing loss are important



Background

The audiogram

Thousands of audiograms are used to conduct these studies.

Background

The audiogram

Thousands of audiograms are used to conduct these studies.

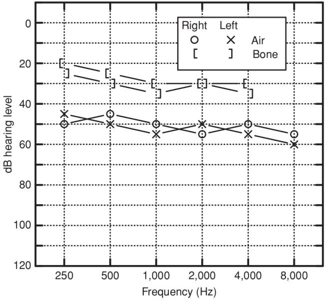


Figure: A typical pure tone audiogram

The Problem

The Problem

Problem 1: Datasets can be polluted with poor quality audiograms

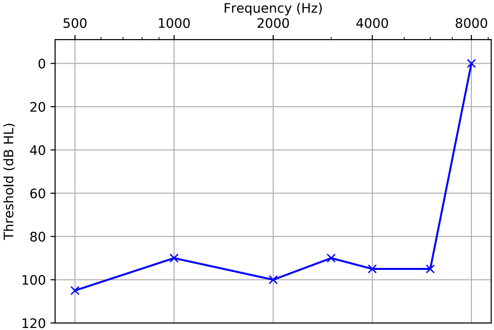


Figure: An audiogram of questionable quality

The Problem

Problem 2: Interpolation-based imputation sometimes fails

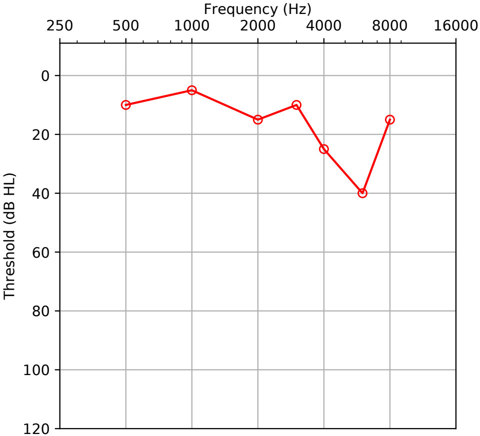


Figure: A case where interpolation fails

The Problem

Problem 2: Interpolation-based imputation sometimes fails

Figure: A case where interpolation fails

The Problem

Problem 2: Interpolation-based imputation sometimes fails

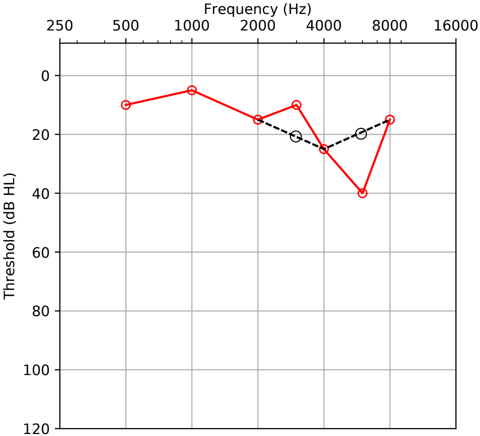


Figure: A case where interpolation fails

The Problem

Research Questions

In our paper, we explore the following questions:

The Problem

Research Questions

In our paper, we explore the following questions:

1. Can we automatically **identify potential quality issues** in large audiograms datasets collected as part of studies on the prevalence of hearing loss?

The Problem

Research Questions

In our paper, we explore the following questions:

1. Can we automatically **identify potential quality issues** in large audiograms datasets collected as part of studies on the prevalence of hearing loss?
2. Can a **data-driven approach** leveraging similarity between audiograms **improve the imputation accuracy** in incomplete audiograms?

Part I: Investigating Unsupervised Learning for Quality Assurance

Part I: Unsupervised Learning for QA

Assumptions

Let us make the following assumptions:

Part I: Unsupervised Learning for QA

Assumptions

Let us make the following assumptions:

1. The anatomy & physiology of the ear limits the possible audiograms

Part I: Unsupervised Learning for QA

Assumptions

Let us make the following assumptions:

1. The anatomy & physiology of the ear limits the possible audiograms
2. The audiograms are manifestations of some fixed, but unknown number of natural processes (diseases of the ear)

Part I: Unsupervised Learning for QA

Assumptions

Let us make the following assumptions:

1. The anatomy & physiology of the ear limits the possible audiograms
2. The audiograms are manifestations of some fixed, but unknown number of natural processes (diseases of the ear)
3. We have access to a large quantity of audiograms from a representative sample of the population

Part I: Unsupervised Learning for QA

Leveraging existing data

In this study, we had access to a large database of 15k+ unlabeled audiograms.

National Health and Nutrition Examination Survey

In this study, we had access to a large database of 15k+ unlabeled audiograms.

nhanes

National Health and Nutrition Examination Survey

How do we leverage this dataset to flag potential quality issues?

Part I: Unsupervised Learning for QA

Modeling the prior probability landscape

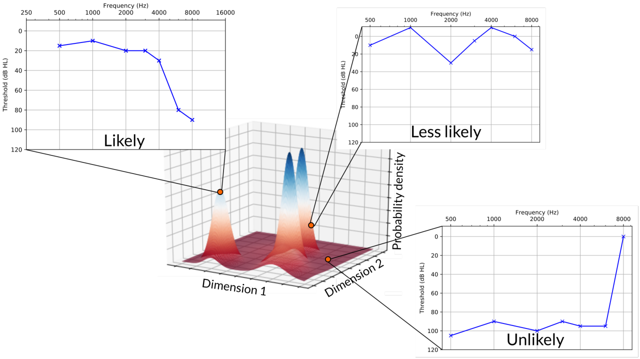


Figure: Conceptual representation of the density estimation problem

Part I: Unsupervised Learning for QA

Modeling the prior probability landscape

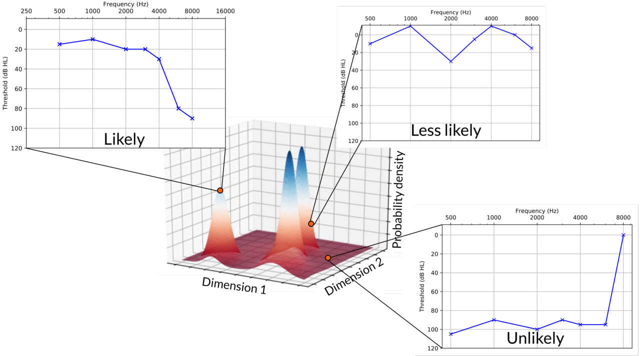


Figure: Conceptual representation of the density estimation problem

That's fine and dandy, but how do we do that?

Part I: Unsupervised Learning for QA

Gaussian Mixture Models

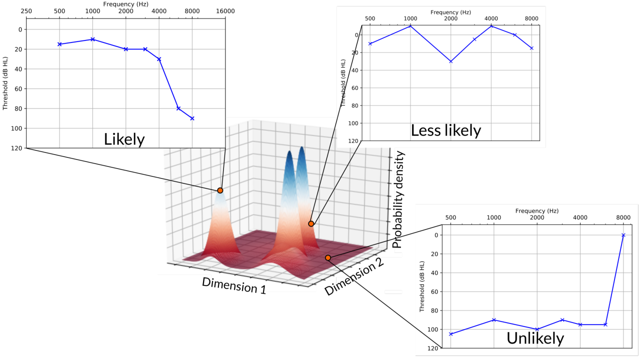


Figure: Conceptual representation of the density estimation problem

$$P(\mathbf{x}) = \sum_{c=1}^{C} \pi_c \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}_c, \mathbf{S}_c)$$

Methodology

Optimizing model complexity

Minimize the Bayesian information criterion (BIC)

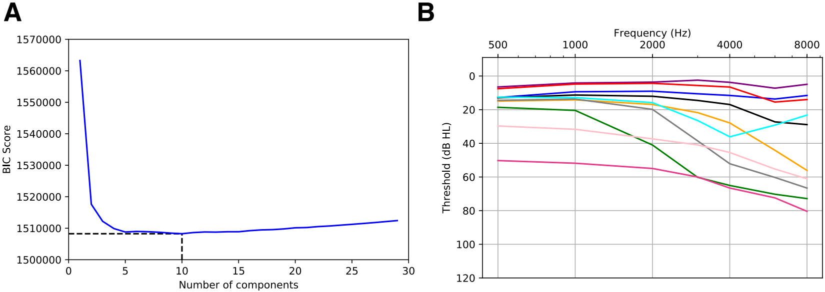


Figure: GMM complexity tuning and resulting components

Part I: Unsupervised Learning for QA

Results

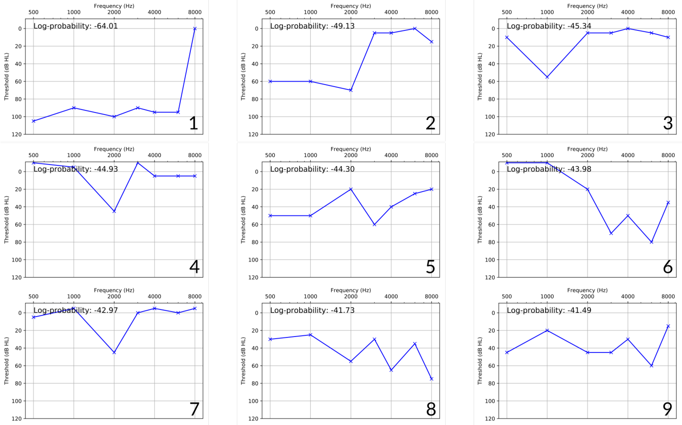
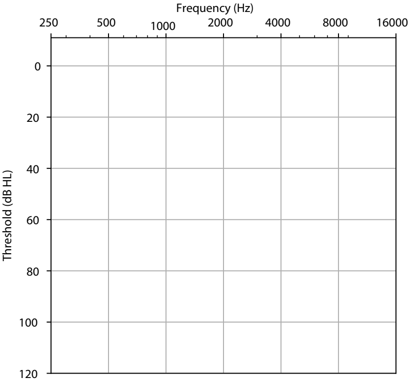


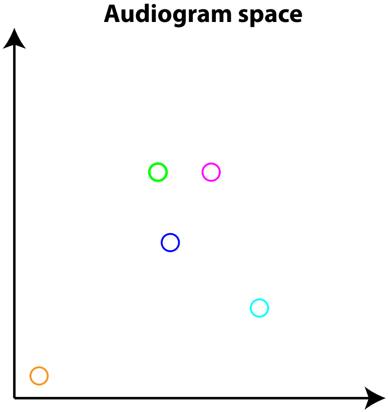
Figure: Audiograms in the NHANES dataset with the lowest log-probability

Part II: Imputing Missing Values

Part II: Imputing Missing Values

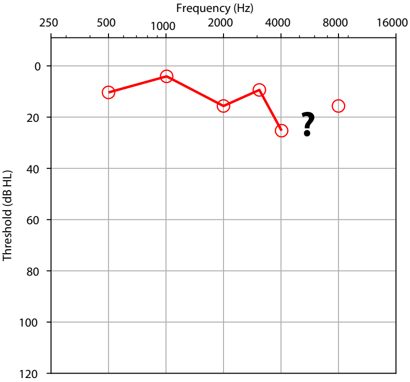
How our k-NN imputation method works

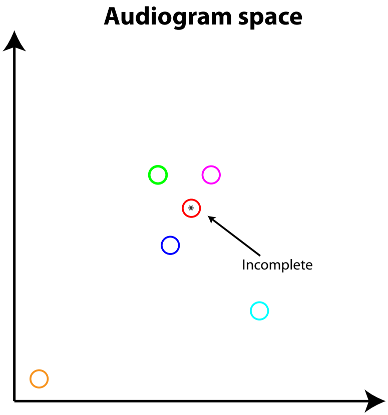




Part II: Imputing Missing Values

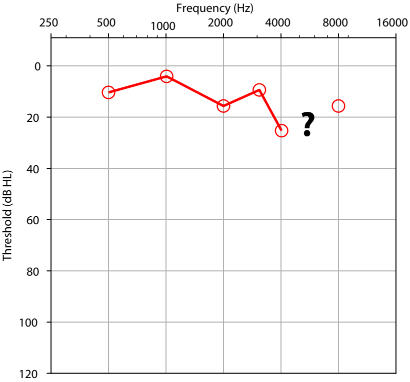
How our k-NN imputation method works

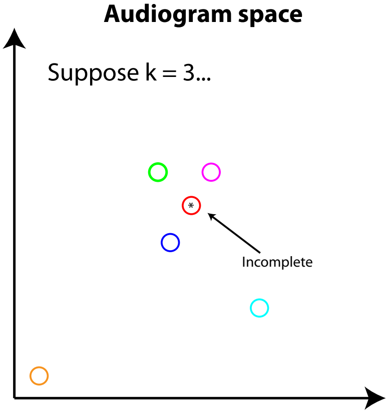




Part II: Imputing Missing Values

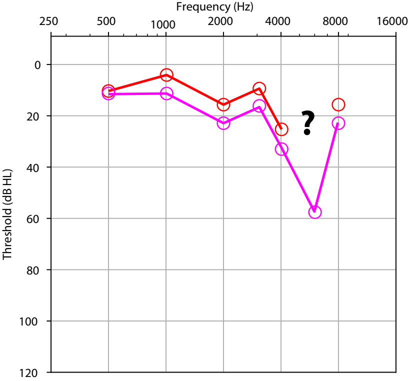
How our k-NN imputation method works

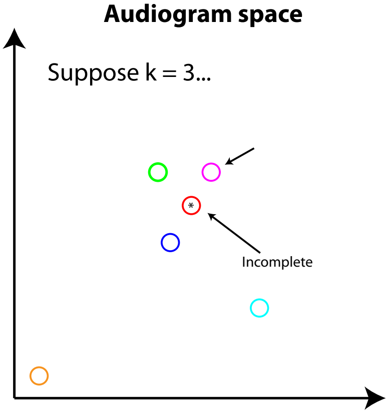




Part II: Imputing Missing Values

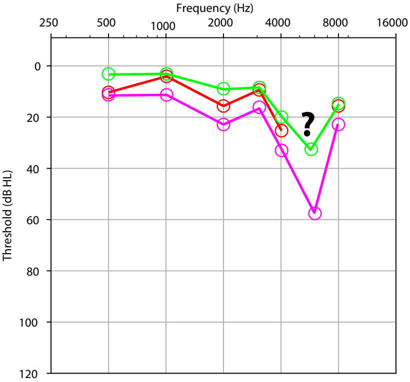
How our k-NN imputation method works

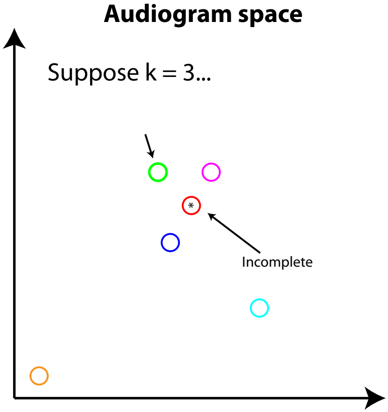




Part II: Imputing Missing Values

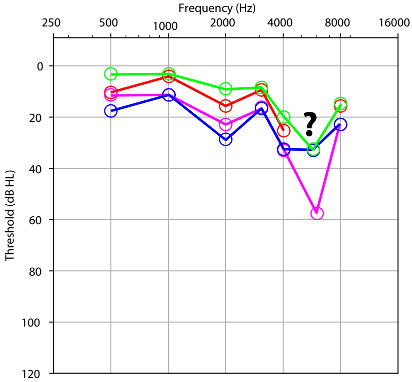
How our k-NN imputation method works

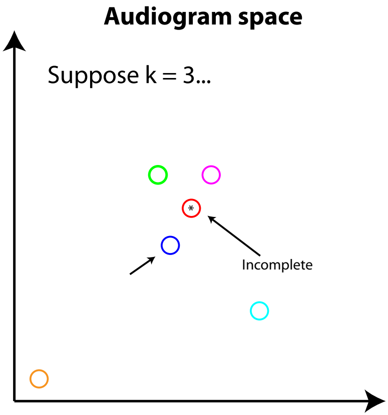




Part II: Imputing Missing Values

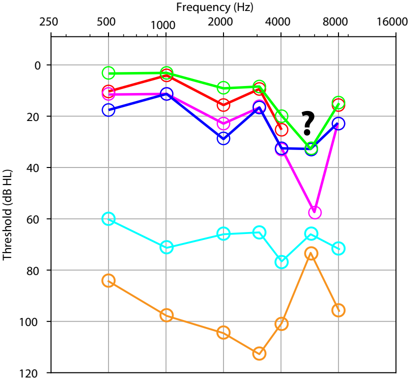
How our k-NN imputation method works

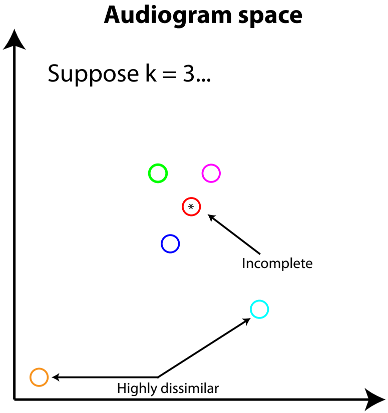




Part II: Imputing Missing Values

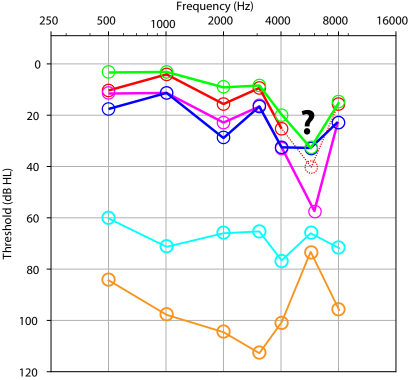
How our k-NN imputation method works

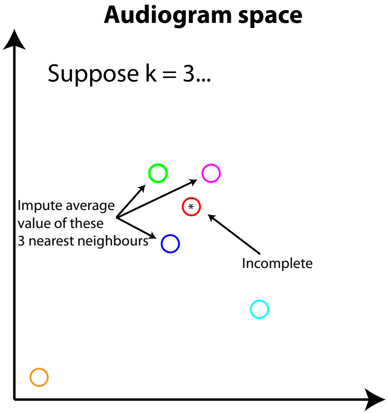




Part II: Imputing Missing Values

How our k-NN imputation method works





Part II: Imputing Missing Values

Our experiment

Adult audiograms

Pediatric audiograms

Training (90% data) **Test (10% data)**

Part II: Imputing Missing Values

Our experiment

nhanes
National Health and Nutrition Examination Survey

Adult audiograms

Pediatric audiograms

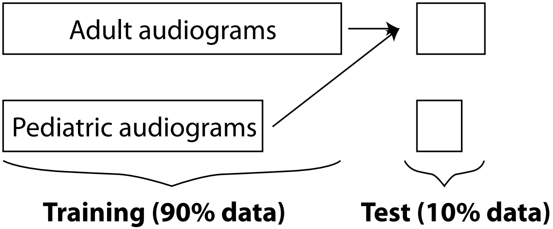
Training (90% data)

Test (10% data)

Part II: Imputing Missing Values

Our experiment

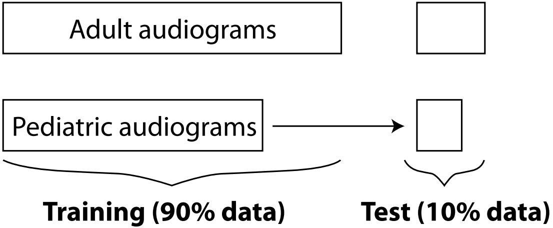
nhanes
National Health and Nutrition Examination Survey



Part II: Imputing Missing Values

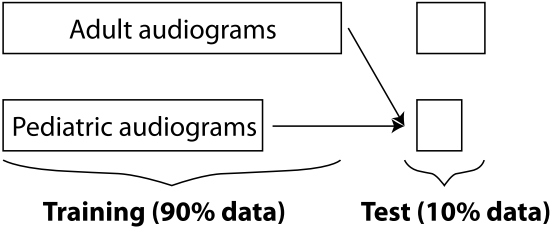
Our experiment

National Health and Nutrition Examination Survey



Part II: Imputing Missing Values

Our experiment



Part II: Imputing Missing Values

Results

Table: Mean absolute error of threshold imputation on a hold-out test set of **pædiatric audiograms**

Training Set	3,000 Hz		6,000 Hz	
	Averaging	k-NN	Averaging	k-NN
All Children	4.68 ± 3.84	4.39 ± 3.54 4.40 ± 3.50	7.66 ± 6.34	6.53 ± 5.19 6.50 ± 5.31

Part II: Imputing Missing Values

Results

Table: Mean absolute error of threshold imputation on a hold-out test set of **pædiatric audiograms**

Training Set	3,000 Hz		6,000 Hz	
	Averaging	k-NN	Averaging	k-NN
All	4.68 ± 3.84	4.39 ± 3.54	7.66 ± 6.34	6.53 ± 5.19
Children		4.40 ± 3.50		6.50 ± 5.31

Table: Mean absolute error of threshold imputation on a hold-out test set of **adult audiograms**

Training Set	3,000 Hz		6,000 Hz	
	Averaging	k-NN	Averaging	k-NN
All	5.49 ± 4.95	5.38 ± 4.75	7.36 ± 6.19	6.92 ± 5.74
Adults		5.39 ± 4.76		6.96 ± 5.74

Closing Statements

Closing Statements

Closing Statements

Conclusions:

- GMM-based approach to QA method is promising

Closing Statements

Conclusions:

- GMM-based approach to QA method is promising
- k-NN approach to threshold imputation dœs no worse than current method

Closing Statements

Conclusions:

- GMM-based approach to QA method is promising
- k-NN approach to threshold imputation dœs no worse than current method

Future work:

- Generalizing our QA method; should identify problematic thresholds locally

Closing Statements

Conclusions:

- GMM-based approach to QA method is promising
- k-NN approach to threshold imputation dœs no worse than current method

Future work:

- Generalizing our QA method; should identify problematic thresholds locally
- Quantitative evaluation of our QA method with trained audiologist(s)

Comments? Questions?