Hypothesis Tests for Continuous Audiometric Threshold Data : Ear and Hearing

- Volume - Issue

- Previous Article

- **Outline**
  - INTRODUCTION
  - STATISTICAL MODELS AND METHODS
  - SIMULATION STUDY
  - CHEARS DATA ANALYSIS
  - DISCUSSION
  - CONCLUSION
  - ACKNOWLEDGMENTS
  - Abbreviations:
  - REFERENCES
  - Supplemental Digital Content
- **Download**
  - PDF
  - EPUB
- **Cite**
  - Copy
  - Export to RIS
  - Export to EndNote
- **Share**
  - Email
  - Facebook
  - Twitter
  - LinkedIn
- **Favorites**
- **Permissions**
- **More**
  - Cite
  - Permissions

Research Article

Hypothesis Tests for Continuous Audiometric Threshold Data

Liu, Zechen[1,6]; Wei, Zhuoran[1,6]; Li, Jiaxuan[1]; Curhan, Gary[2,3,4,5]; Curhan, Sharon[2,3,4,7]; Wang, Molin[1,2,3,4,7]

Author Information

[1]Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, USA;

[2]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, USA;

[3]Harvard Medical School, Boston, Massachusetts, USA;

[4]Channing Division of Network Medicine, Department of Medicine, Brigham and Women’s Hospital, Boston, Massachusetts, USA; and

[5]Renal Division, Department of Medicine, Brigham and Women’s Hospital, Boston, Massachusetts, USA;

[6]These authors contributed equally as co-first authors;

[7]These authors contributed contributed equally as co-senior authors.

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and text of this article on the journal’s Web site (www.ear-hearing.com).

Received January 25, 2023; accepted February 23, 2024

Address for correspondence: Molin Wang, Departments of Epidemiology and Biostatistics, Harvard T.H. Chan School of Public Health, 665 Huntington Ave, Boston, MA 02115, USA. E-mail: mwang@hsph.harvard.edu

Ear & Hearing ():10.1097/AUD.0000000000001503, March 28, 2024. | DOI: 10.1097/AUD.0000000000001503

- Open
- SDC
- PAP

Metrics

Abstract

Objectives:

Hypothesis tests for hearing threshold data may be challenging due to the special structure of the response variable, which consists of the measurements from the participant’s two ears at multiple frequencies. The commonly-used methods may have inflated type I error rates for the global test that examines whether exposure-hearing threshold associations exist in at least one of the frequencies. We propose using both-ear methods, including all frequencies in the same model for hypothesis testing.

Design:

We compared the both-ear method to commonly used single-ear methods, such as the worse-ear, better-ear, left/right-ear, average-ear methods, and both-ear methods that evaluate individual audiometric frequencies in separate models, through both theoretical consideration and a simulation study. Differences between the methods were illustrated using hypothesis tests for the associations between the Dietary Approaches to Stop Hypertension adherence score and 3-year change in hearing thresholds among participants in the Conservation of Hearing Study.

Results:

We found that (1) in the absence of ear-level confounders, the better-ear, worse-ear and left/right-ear methods have less power for frequency-specific tests and for the global test; (2) in the presence of ear-level confounders, the better-ear and worse-ear methods are invalid, and the left/right-ear and average-ear methods have less power, with the power loss in the left/right-ear much greater than the average-ear method, for frequency-specific tests and for the global test; and (3) the both-ear method with separate analyses for individual frequencies is invalid for the global test.

Conclusions:

For hypothesis testing to evaluate whether there are significant associations between an exposure of interest and audiometric hearing threshold measurements, the both-ear method that includes all frequencies in the same model is the recommended analytic approach.

INTRODUCTION

Hearing loss was a leading cause of disability in 2015; over 5% of the world's population suffers from disabling hearing loss (Wilson et al. 2017; Chadha et al. 2021; WHO 2021). Moreover, the adverse influence of hearing loss on health and quality of life is considerable (Mick et al. 2014; Dawes et al. 2015), thus research to identify potentially modifiable risk factors for hearing loss that could inform strategies for prevention is a pressing public health priority (Chadha et al. 2021; Haile et al. 2021).

In epidemiological studies of hearing health, researchers are often interested in assessing whether associations exist between a given exposure and the hearing threshold measures. However, statistical methods for the evaluation of exposure-hearing associations using audiometric threshold data are inconsistent (Cruickshanks et al. 2003; Bainbridge et al. 2008). Researchers commonly used a single-ear (e.g., the worse-ear [WE], better-ear [BE], left/right-ear) hearing measurement as the response variable (Verschuur et al. 2012; Grondin et al. 2015; Lin et al. 2016) or analyze the data for each frequency or group of frequencies separately (Hu et al. 2018; Shih et al. 2020). Definitions of the single-ear method may vary and are typically based on the pure-tone average (PTA) of threshold measurements at a prespecified set of frequencies. In a longitudinal setting, the outcome could be the frequency-specific PTA threshold measurements at each follow-up time point, or alternatively the change in frequency-specific PTA at each postbaseline follow-up time point. The WE method and BE method are based on measurements of the WE, the ear with a higher threshold or change in threshold, and the BE, the ear with a lower threshold or change in threshold, respectively. Table 1 shows an example of the WE and BE outcomes for the low-frequency PTA threshold measurements. In the longitudinal study setting, the ear which is the WE or BE could change over time. For example, in Table 1, the WE is the left-ear at visit 1 but the right-ear at visit 2.

TABLE 1. - Example of PTA of threshold data for low frequencies (0.5, 1, 2 kHz) from 1 participant in a hearing loss study
Scroll left or right to view entire table.

Visit	Single-Ear Method					Both-Ear Outcome
	Right	Left	WE Outcome	BE Outcome	AE Outcome	
Visit 1 (baseline)	15	10	15	10	12.5	(15, 10)
Visit 2	15	20	20	15	17.5	(15, 20)
Change	0	10	10	0	5	(0, 10)

Data in this table are fabricated.
AE, average-ear; BE, better-ear; Left, left-ear; PTA, pure-tone average; Right, right-ear; WE, worse-ear.

Existing publications demonstrate the performance of different estimation methods (e.g., bias and efficiency of WE/BE, average-ear [AE], and both-ear methods) for continuous outcomes (Sheng et al. 2022) and binary outcomes (Chen et al. 2022). This article focuses on continuous outcomes. For continuous outcome scenarios, when there are only participant-level confounders, using WE or BE methods leads to unbiased but less efficient estimators; lower efficiency means the estimators have larger variance (Sheng et al. 2022). If the information from only one ear at each time point is used in the analyses, then the information from the other ear is ignored. Therefore, the AE method which uses the threshold measures averaged over the two ears as the outcome, and the both-ear method, which uses the threshold measures from both of the two ears, may be preferred (Sheng et al. 2022). See Table 1 for an example of AE and both-ear outcomes. Note that, for presentational simplicity and to distinguish from the both-ear method which uses both ears' data as a cluster, we categorize AE as a single-ear method throughout this article even though its outcome is derived from both ears' data. In the presence of ear-level confounders, failure to include them in the model would lead to biased estimates for all methods. Previous studies demonstrate that if the ear-level confounders are included in the model, then the WE and BE methods may still lead to biased estimators. However, in this case, the bias would be less than that in models where the ear-level confounders were not included. Moreover, the AE method may lead to unbiased, but less efficient, estimators compared with the both-ear method (Sheng et al. 2022). On the other hand, biased estimators do not necessarily imply invalid hypothesis tests (Stahlecker & Schmidt 1996). A valid test means that the type I error rate, the probability of rejecting the null hypothesis when the null hypothesis is true, is controlled under the significance level, which is typically set at 5%. A test is invalid if the type I error rate is greater than the significance level. Similarly, estimators with higher efficiency do not necessarily imply hypothesis tests with higher power (Sundrum 1954). A powerful test means that the type II error rate, the probability of not rejecting the null hypothesis when the alternative hypothesis is true, is small. More specifically in terms of mathematical formula, statistical power = (1 – type II error). In commonly used hearing threshold data analysis methods, PTA measurements based on different groups of frequencies (e.g., low-frequency PTA, mid-frequency PTA, high-frequency PTA) are analyzed separately in different models instead of simultaneously in the same model. Thus, if one concludes that there is an exposure-outcome association because at least one of the tests in the separate analyses for individual frequency groups is significant, then issues with multiple comparisons and the correlations between the estimated parameters for these frequency groups are ignored, leading to an inflated type I error. In this case, one can use adjustment strategies to account for multiple comparisons, such as the Bonferroni method. However, these adjustment methods may not be able to easily address highly correlated hypotheses (Chen et al. 2017).

Previous studies that compared the single-ear and both-ear methods focused only on estimators for the exposure-hearing association (Chen et al. 2022; Sheng et al. 2022), while less concerned about hypothesis testing. In this study, our objective is to evaluate the validity and power of the statistical tests to evaluate exposure-hearing associations using both-ear and single-ear methods. We consider two methods for hypothesis tests based on both-ear data: (1) methods based on regression analyses inclusive of all frequencies in the same model, and (2) methods based on separate regression analyses where each model evaluates a single frequency. For the single-ear methods, we consider WE, BE, left/right-ear method, and AE; all frequencies are included in the same model for all the single-ear methods. We will consider both hypothesis tests for each frequency and a global test across all frequencies. The generalized estimating equation (GEE) method is used to obtain estimates and their variances for the regression coefficients of the regression models, taking into account the within-participant correlation between hearing threshold measurements in both ears and across frequencies, if applicable.

STATISTICAL MODELS AND METHODS

When only one frequency group (i.e., low-frequency PTA, mid-frequency PTA, high-frequency PTA) is included in the analysis, we can assume the following both-ear model:

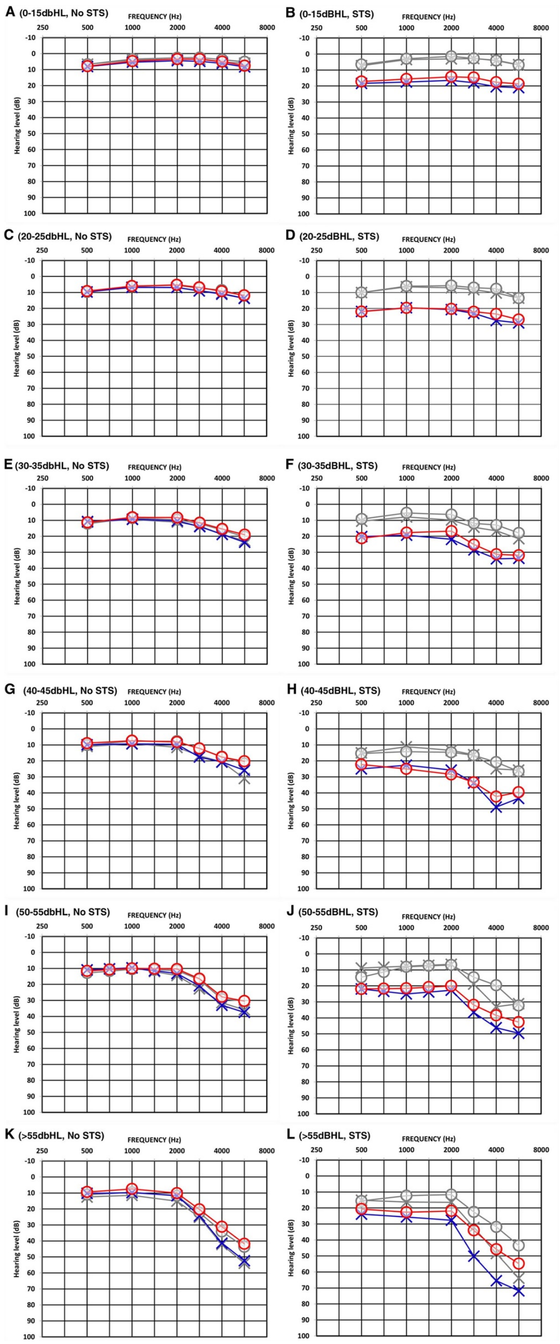
$$E\left(Y_{i,j}|X_i,\ W_i,\ Z_{i,j}\right)=\beta_0+\beta_1X_i+\beta_2W_i+\beta_3Z_{i,j}$$

where i indexes participants ($i=1,2,3,\ldots,N$), j is the index for the ear ($j=0,1$), and $Y_{i,j}$ is the hearing threshold measurement for individual i and ear j. Among the independent variables, X represents the participant-level exposure of interest, W is possibly vector-valued participant-level potential confounders, and Z represents possibly vector-valued ear-specific potential confounders (e.g., baseline hearing thresholds). Without further specification, parameters are row vectors and variables are column vectors throughout this article.

If multiple frequencies are considered, we need to add covariate-frequency interactions in the both-ear model, leading to

$$E\left(Y_{i,j,q}|X_i,\ W_i,\ Z_{i,j}\right)=\beta_{0q}+\beta_{1q}X_i+\beta_{2q}W_i+\beta_{3q}Z_{i,j}$$

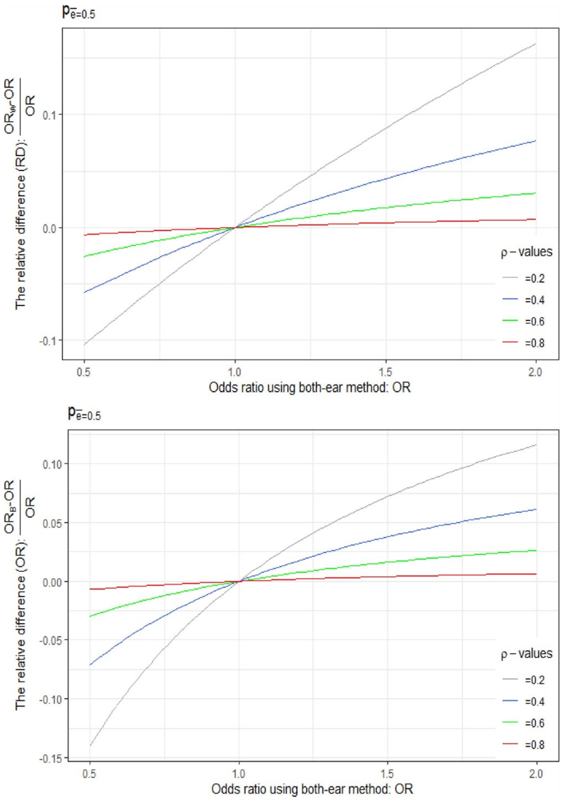
where $E\left(\cdot\right)$ stands for the expected value, $q\ \ (q=1,\ldots Q)$ indexes the pure-tone frequency, β_{1q} represents the association of the exposure with the hearing threshold level under



-

Pre-Existing Audiometric Hearing Loss is a Predictor of Significant Threshold Shift Following Injury During Combat Deployment

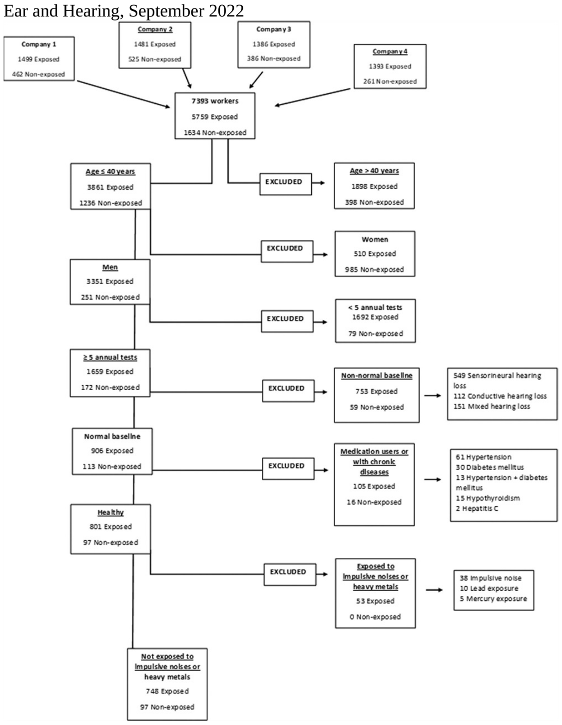
Ear and Hearing, September 2023



-

Both-Ear Method for the Analysis of Audiometric Data

Ear and Hearing, September 2022



-

Five-Year Longitudinal Cohort Study Determines the Critical Intervals for Periodic Audiometric Testing Based on 5070 Tests of Metallurgical Workers Exposed and Nonexposed to Noise

Ear and Hearing, January 2022