Original Research Article

Toward a test battery for differential categorization of age-related hearing loss

Rachael R. Baiduc[1], Jungwha Lee[2], Sumitrajit Dhar[1,3]

[1]Department of Communication Sciences and Disorders, Northwestern University, Evanston, [2]Department of Preventive Medicine, Biostatistics Collaboration Center, Northwestern University, Chicago, [3]Knowles Hearing Center, Roxelyn and Richard Pepper, Northwestern University, Evanston, Illinois, USA

Abstract

Age-related hearing loss (ARHL, or presbycusis) results from neural and/or cochlear degeneration. A taxonomy distinguishing presbycusis subtypes according to site of lesion was originally proposed by linking audiometric results to histopathological findings. In most cases, the pathology is complex and audiometry and word recognition scores (WRS) are insufficient to characterize pathologies of the auditory periphery. Several sophisticated tests of auditory function, with some specifically designed to inspect cochlear or neural status (e.g., distortion product otoacoustic emissions [DPOAEs] and auditory brainstem response [ABR]) are available but not in routine use to distinguish between presbycutic subtypes. There are no *in vivo* methods in place to identify contributing pathologies and their relative dominance in individual instances of presbycusis. However, the promise of upcoming therapies (genetic, pharmaceutical, etc.) cannot be realized without accurate identification of presbycusis subtypes. The goal of the present study was to investigate possible improvements in differential categorization of presbycutic subtypes. We explored a test battery composed of behavioral (audiometry and speech testing) and physiological (ABR, DPOAEs, and electrocochleography) assays in presbycutic ears to ask if improvements beyond the "gold standard" (behavioral thresholds through 8 kHz and word recognition) are possible. Data from 10 hearing impaired (HI) individuals were compared to those from 21 normal hearing (NH) adults. Exploratory factor and hierarchical cluster analyses (EFA and HCA respectively) were used to evaluate phenotyping strategies. The EFA revealed three factors (highest audible frequency (HAF), pure-tone average (PTA), and $2f_1$-$f_{2(High)}$ DPOAEs) that accounted for most of the variability in hearing outcomes among the 31 participants. Hierarchical cluster analysis using the gold standard and enhanced multivariate approach revealed: (1) The clinical gold standard distinguished NH and HI participants, but failed to find commonalities among individuals with similar hearing profiles and (2) The enhanced test battery grouped participants with similar profiles, presumably indicating an underlying relationship in pathophysiology. Model data support the feasibility of a finer-grained categorization of presbycusis than is available in current practice, although more data are needed to understand the complexities of phenotyping.

Key words: Aging, distortion product otoacoustic emissions, electrophysiology

Access this article online	
Quick Response Code:	Website: www.jisha.org
	DOI: 10.4103/0974-2131.162426

Introduction

Hearing loss is a highly prevalent chronic health condition, with 360 million individuals affected worldwide and a predicted burden in excess of 44 million by 2030 in the United States alone.[1] These prevalence estimates encompass both congenital and acquired cases. The disease can substantially impact

Address for correspondence:
Dr. Rachael R. Baiduc, Searle Center for Advancing Learning and Teaching, Northwestern University, 627 Dartmouth Place, 111, Evanston, Illinois, 60208, United States. E-mail: RRBaiduc@northwestern.edu

quality of life, causing psychosocial and financial difficulties, particularly for the elderly.[2] Auditory aging, or presbycusis, begins early in life and to some extent, may be unavoidable in industrialized societies. Even individuals aged 36–45 years have slightly higher (poorer) behavioral hearing thresholds above 10 kHz than those between 10–21 years.[3] Data also suggest significant individual variability in trajectories of auditory aging. The exact pathophysiology related to these differences is unknown at this time.

Peripheral presbycusis can originate from age-related degeneration of cells and structures at various points along the peripheral auditory pathway including the stria vascularis, the site of endolymph production and biological battery of the cochlea (metabolic or strial subtype),[4,5] the outer hair cells (sensory),[6] the auditory nervous system, specifically fibers of the spiral ganglion (neural),[7] and supporting structures (mechanical).[8] Taken together, these points of cellular degeneration are known as 'sites of lesion' and to date, are considered indistinguishable using today's regular clinical tests which patients are likely to undergo in a routine assessment in most audiology clinics.

A taxonomy distinguishing presbycusis according to site of lesion was originally proposed by Schuknecht and colleagues[4,6,8-11] who built upon the seminal work of Crowe *et al*. (1934)[12] by linking audiometric, word recognition (when available), and histopathological findings from human temporal bones.[6,8,11] Originally, Schuknecht[8,10,11,13] segregated presbycusis into sensory (hair cell), strial, and neural subtypes based largely on audiometric shape, or phenotype, and corroborative pathophysiological findings. Since these seminal studies, there have been numerous other attempts to correlate site of lesion and behavioral, physiological, or histopathological findings to various degrees of success.[14-18] For example, Nelson and Hinojosa[17] evaluated histopathology in six individuals with flat audiograms. Examination of these audiometric phenotypes alone would suggest all six cases had strial presbycusis, but using a novel approach to measure strial volume, they determined only one of the six had significant strial degeneration. Taken together, appraisal of these studies indicates the traditional approach using the audiogram is insufficient to differentiate types of presbycusis, in part because of the overlap in audiometric phenotypes caused by multiple pathologies (e.g., noise exposure, ototoxic drugs, cardiovascular disease, etc.). Age-related hearing loss (ARHL) affects multiple cellular entities to varying degrees, but is currently evaluated with rather rudimentary tools such as audiometry and word recognition (the clinical "gold standard") in the majority of settings where patients receive hearing healthcare. These tests are widely applied and cost-effective but they assess functional aspects of the *entire* auditory pathway and therefore are not well equipped for differential diagnosis of peripheral pathologies. At the time of the original presbycusis phenotyping studies, many of the extensive behavioral and physiological assays of the present day were not available. Today, there are numerous ways to appraise auditory function and inspect outer hair cell and neural status non-invasively and objectively (e.g., distortion product otoacoustic emissions [DPOAEs] and auditory brainstem response [ABR] respectively). Electrocochleography (EcochG) may also be useful because it has been employed clinically for years to recognize endolymphatic hydrops, a specific type of strial pathology, and by extension, may be valuable for characterization of strial presbycusis. Despite the widespread availability of these physiological tests there remains no *in vivo* method to differentiate various forms of presbycusis. It is our view that presbycusis is a complex disease; a juxtaposition of inherent aging processes and environmental, or external, factors. In this pilot effort to differentially categorize presbycusis, we thus consider the representation of hearing loss in individuals who also present with other risk factors for hearing loss (e.g., noise exposure, cardiovascular disease, etc).

The goal of differential diagnosis is straightforward: provide patients with individualized care, especially in regards to burgeoning therapeutics targeting specific cellular dysfunctions that will one day become available in the clinic. Hair cells and auditory neurons are incapable of spontaneous regeneration (without intervention). With the recent demonstration of mammalian hair cell regeneration using a gamma-secretase inhibitor,[19] preventive and therapeutic strategies are of increasing interest to the research and clinical communities. Prophylactic agents that prevent spiral ganglion or hair cell death as well as those that promote re-growth are being explored.[20-27] Gene or stem cell therapeutics to regenerate hair cells are also emerging.[21,25,27] The current evaluation strategy (the gold standard) may be sufficient for hearing aid selection and fitting, but precise diagnostic techniques to pinpoint dominant site(s) of lesion will be required to determine candidacy and dosing for these up-and-coming therapies.

We conjecture that audiometry and word recognition alone are not capable of parsing presbycusis subtypes, thus we employed a sophisticated battery comprised of both behavioral (audiometry from 0.125–20 kHz and speech testing in quiet and noise) and physiological

(DPOAE, ABR, and EcochG) tests to begin exploring a classification scheme for peripheral presbycusis. We aim to identify combinations of tests that can segregate presbycutic subtypes in a way superior to the standard approach. The chosen tests evaluate various (and presumably non-overlapping) aspects of auditory function hence they are well situated to address this aim. In this report, we compare responses on these tests between 21 normal hearing (NH) young adults who serve as a reference group and 10 case participants with hearing loss (nine with presbycusis and one young adult with NH and trouble hearing in noise [NH: THIN]).

The work in this paper reflects an initial attempt to develop a scheme for differential classification of presbycusis in living patients. There are numerous gaps in the literature regarding the complexity of presbycusis, especially in regards to physiological findings and their relationship to the audiometric phenotype. Here we attempt to fill one of those voids. We hypothesize that a multivariate test battery will provide finer-grained identification of presbycusis subtypes compared to behavioral hearing thresholds (through 8 kHz) in conjunction with word recognition (the clinical "gold standard"). To this end, we approach our dataset both qualitatively and quantitatively in an exploratory fashion. Our comprehensive data provide an extensive physiological and behavioral phenotyping scheme in NH and hearing impaired (HI) individuals. Ultimately, we determine whether or not any test(s) in our series can provide information more salient for differential categorization of presbycusis than the standard method utilizing audiometric phenotype and speech testing. We also examine the facility of a multivariate test battery to cluster similar cases of presbycusis. This work is undoubtedly still in its infancy but lays a scaffold which can be built upon by future studies.

Materials and Methods

Selection and description of the normal hearing/ control participants

Twenty-one NH adults (5 male, 16 female, mean age = 23.52 years, SD = 4.09) were recruited for this study. All participants underwent a case history evaluation and individuals with history of noise exposure, self-reported difficulty hearing in quiet or noisy situations greater than 3 on a 1–10 scale, medical conditions associated with hearing loss (e.g., cardiovascular disease, diabetes, cancer, tinnitus, dizziness, etc.), current/previous use of ototoxic medications, exposure to solvents, or history of cigarette smoking were excluded from the NH group. Participants were required to pass an otoscopic examination (ear canals visibly free of excessive debris, permitting clear visualization of the tympanic membrane) as well as audiometry (pure tone thresholds ≤20 dB HL from 0.25–8 kHz) and tympanometry (*details below*).

Participants with hearing impairment

Nine adults (three female) with sensory/neural hearing loss (defined as one or more behavioral thresholds ≥20 dB HL from 0.25–8 kHz) in one or both ears were selected for participation in the HI (case) group. In addition, one young individual (24-year-old female) with clinically normal hearing through 8 kHz but severe difficulty hearing in noise was included, yielding a total of 10 HI participants. Although we cannot be certain of the exact pathology in this individual, her inclusion was a means to test the capacity of the statistical approaches to appropriately identify such individuals as 'normal' and served as a proof of concept that the models were operating accurately. In addition, we hypothesized that a multivariate test battery might point toward underlying pathology in a way the gold standard battery could not. Participants in the HI group ranged in age from 24–70 years (mean age = 56.80 years, SD = 11.89). They were required to pass otoscopic and tympanometric screenings.

Technical information: Measurements and equipment

All measurements were conducted in a double-walled sound attenuated test chamber with participants comfortably seated in a reclining chair. Informed consent was obtained prior to enrollment in the study. Participants were compensated for their involvement. The study protocol and procedures were conducted in compliance with and approved by the Institutional Review Board of the University. The study was performed in accordance with the ethical standards outlined by the Declaration of Helsinki (amended 2013).

Customized software (developed by C. Talmadge)[28] was employed for sound source calibration, signal generation (threshold tracking and DPOAEs), and emission recordings. A MOTU828 MKII Firewire device was used for analog-to-digital and digital-to-analog conversion (sampling rate of 44.1 kHz, 24 bit) for DPOAEs and Békésy audiometry. Signals were generated, delivered to the MOTU for digital-to-analog conversion, to an Etymotic Research H4C low distortion power amplifier, and to MB Quart 13.01HX transducers. Speakers were coupled to an Etymotic Research ER-10B+ probe assembly and DPOAEs were recorded using the probe apparatus and preamplifier (+20 dB gain). Emission recordings were digitized by the MOTU and stored on the computer for later analysis.

The complex geometry of the human ear canal can lead to significantly different stimulus levels at the tympanic membrane at frequencies above 6–8 kHz when calibrated using standard means such as an artificial ear. Because measurements were made in this study above 8 kHz, a different approach was needed. Details of the calibration procedure have been published elsewhere.[3] Ear-specific *in situ* depth calibration was carried out at the start of each experimental session by playing a chirp to the test ear, estimating insertion depth in the ear canal, and applying correction factors for that insertion (see Lee *et al*., 2012 for details).[3]

Behavioral assessments

Participants underwent a range of threshold and supra-threshold auditory tests. First, audiometry was performed on an Interacoustics Audio Traveller AA220 with ER-3A insert earphones using the Hughson-Westlake method[29] to ascertain thresholds between 0.25–8 kHz. All participants in the NH group had pure tone thresholds ≤20 dB HL at every test frequency (criteria for 'pass'). The ear with the better thresholds was designated as the test ear for those in the NH group. For the 10 participants with HI, the shape of the audiogram was visually examined and an audiometric classification was assigned (e.g., sloping, notched, etc.) as is done in routine clinical practice. Specific criteria for these classifications was not employed at this stage of the study, although in general, we described audiograms with poorer high frequency thresholds than low frequency as sloping, those with normal mid-frequency hearing and poor low and high frequency thresholds as inverted cookie bite, those with a discernable notch >10 dB at 2 or 4 kHz as notched, and those with minimal variation across frequency as flat. For the HI group, test ear selection was based on severity of loss (mild to moderate losses were necessary) and ear canal characteristics, keeping in mind the desire to recruit cases with a range of audiometric profiles. Tympanometry was performed with the Interacoustics Audio Traveller AA220 and test ears demonstrated normal immittance findings in accordance with clinical standards.[30] Table 1 presents a summary of the HI participants including age, gender, test ear, audiometric configuration, pure-tone average (PTA) in the test ear (mean threshold in dB HL at 0.5, 1, and 2 kHz), and notable medical items from the case history.

The remaining behavioral and physiological tests were carried out only in the test ear. A modified Békésy tracking procedure[31] was used to determine thresholds at 21 frequencies from 0.125–20 kHz from which we identified the highest audible frequency (HAF; the highest frequency that elicited a threshold response). Pulsed tones (250 msec, 25 msec rise/fall time) were presented twice per second at the test frequency. Details of the procedure can be found in Lee *et al*. (2012).[3]

Speech testing was performed on 19 NH participants and all 10 individuals with HI. Custom software (developed by C. Chan[3]) was used for speech reception threshold (SRT), word recognition score (WRS), and the Words in Noise (WIN) test. All stimuli were taken from recordings and digitized to permit delivery through our software interface. The MB Quart 13.01HX transducer was connected to an ER3-14A foam ear tip with plastic tubing to deliver speech stimuli. The SRT and WRS were determined using standard clinical protocol, with stimuli for WRS presented at 40 dB SL re: SRT. Speech understanding in noise performance was assessed using the WIN test on 18 of the NH participants and all 10 participants with HI in accordance with the developer's guidelines.[32,33] The WIN test uses monosyllabic NU-6 words spoken by a female speaker in a background of multi-talker babble. The level of the background speech was fixed at 70 dB SPL for the NH group (individuals with a PTA ≤25 dB HL). For participants with PTAs between 25–40 dB HL, a presentation level of 80 dB SPL was used. Five words were presented at each of seven unique SNR conditions from 0–24 (in 4 dB steps).

Physiological testing

DPOAEs ($2f_1$-f_2 and f_2-f_1) were recorded using an f_2/f_1 ratio of 1.14 and stimulus intensity of $L_1 = L_2 = 70$ dB SPL. These stimulus parameters were chosen to maximize f_2-f_1 amplitude based on pilot data from our laboratory (*data not shown*). Tones were swept logarithmically at 8 s/octave between ~0.7 and 7 kHz (f_2) and at 24 s/octave thereafter. At least six sweeps were recorded with 2 s intervals between them. The higher frequency primary (f_2) was swept from approximately 0.664–19.221 kHz, resulting in DP frequencies between 0.08 and 2.4 kHz for f_2-f_1 and 0.5 and 14.5 kHz for $2f_1$-f_2. Recordings were processed offline using a least squares fit (LSF) algorithm to generate DPOAE level estimates.[28,34,35] The analysis window was 22050 points (or 0.5 seconds), with 90% overlap between neighboring windows and 63 initial points skipped. DPOAE level was estimated every 1–4 Hz at low frequencies (0.08–1.0 kHz) and every 5–13 Hz for higher frequencies (1.2–14.5 kHz).

Data were smoothed and binned prior to graphing and statistical analyses. Smoothing was done in 50-point bins with 90% overlap between adjacent bins. The mean of all points within a third-octave range centered at specific DP frequencies was used to calculate the nominal DPOAE level for that frequency.

Table 1: Demographic, audiometric, and medical history data for HI participants. Noise exposure indicates self-reported occupational or recreational exposures

Case number	Sex	Age (years)	Test ear	Audiometric configuration	Pure-tone average (dB HL)	Notable medical history
1	M	57	L	Sloping	20.00	Noise exposure (both), tinnitus, vertigo, HPN, former smoker
2	M	64	L	Sloping	20.00	Noise exposure (occupational), tinnitus
3	M	62	R	Sloping	18.33	Slight tinnitus
4	M	56	L	Sloping	16.67	Heart attack
5	M	55	L	Sloping	18.33	Noise exposure (occupational)
6*	M	55	R	Flat	16.67	see case 5
7	M	66	R	Notch	11.67	Noise exposure (recreational), former smoker
8	F	70	L	Inv. Cookie	18.33	Former smoker, HPN, slight tinnitus
9	F	59	R	Inv. Cookie	30.00	Tinnitus, vertigo, otalgia
10	F	24	L	NH: THIN	6.67	ADHD

ADHD: Attention deficient hyperactivity disorder (self-reported); HPN: Hypertension (and/or treatment with blood pressure lowering drugs); Inv. Cookie: Inverted cookie bite audiogram (low and high frequency loss); NH: THIN: Normal hearing with trouble hearing in noise; M: Male; F: Female; L: Left; R: Right; *Participants 5 and 6 are the same 55 year-old male, with separate testing performed for each ear

Center frequencies varied between the two DP types ($2f_1$-f_2 and f_2-f_1). That is, f_2-f_1 data were binned around DP frequencies spanning 0.5–2 kHz and $2f_1$-f_2 data around 0.5–12.5 kHz. We display all the smoothed/binned data and recommend that caution should be exercised when interpreting these preliminary results as no signal-to-noise ratio (SNR) criteria have been applied to preselect data. Preselecting using SNR criteria was not done as standards for the f_2-f_1 DPOAE are not yet established. Because the noise floor was typically higher for f_2-f_1 than $2f_1$-f_2, the SNR of the DPOAE responses for each emission type (cubic and quadratic) was considered in the analytical models instead of the DPOAE level. Responses were averaged into a low (≤8kHz) and high (>8 kHz) frequency group based on f_2. Statistical analyses of DPOAEs thus included four groups ($2f_1$-$f_{2(Low)}$, $2f_1$-$f_{2(High)}$, f_2-$f_{1(Low)}$, f_2-$f_{1(High)}$).

For electrophysiological testing (ABR and EcochG) and analyses, a Biologic Auditory Evoked Potentials (AEP) instrument (version 6.2.0) connected to an HP laptop computer was used. A single-channel vertical montage was employed for the ABR measurements, with Natus Silver-Silver Chlorided 10 mm disc electrodes attached with 1.0 M silicone lead wires adhered to the skin of the forehead (F_z; inverting electrode), opposite mastoid (ground) and test mastoid (non-inverting electrode). Prior to recording, the impedances of the three electrode sites were checked for target values of <5 kΩ.

Brainstem responses were elicited using rarefaction clicks (17.7/sec) delivered monaurally to the test ear using Biologic insert earphones with ER3-14A foam ear tips or E-A-RLink 3A/5A tips for small canals. The initial presentation level was 80 dB nHL and was lowered in 10 dB steps (5 dB closer to threshold) until wave V was no longer identifiable. Two or three trials of 1000 repetitions were carried out at each level to ensure repeatability, resulting in 2000–3000 click responses per stimulus condition. Sweeps were comprised of 512 digitized points. Artifacts greater than 23.80 μV were rejected online and not considered in the average response for that stimulus level. A recording window (epoch) of 10.66 msec was employed with a 1.29 pre-stimulus period. Responses were amplified (×100,000) and filtered digitally between 30–1500 Hz. After averaging two or more repetitions, ABR waveforms were examined visually. The absolute latency (in msec) and amplitude (in μV) of wave V was determined. The lowest intensity stimulus (in dB nHL) that elicited a repeatable wave V was deemed threshold. Although latencies were also recorded, because this report was limited in the number of variables (test outcomes) that could be considered statistically, we chose to include ABR wave V threshold as the only ABR variable incorporated in the statistical models. As ABR threshold reflects both the number of surviving spiral ganglion neurons as well as neural synchrony (see Boettcher, 2002[36] for review), it is an appropriate starting place for our analytical approach as it provides a means to identify reduction in the spiral ganglion fiber population.

Extra-tympanic EcochG was performed on 18 NH individuals and all 10 HI participants. Single-channel recordings were made using a horizontal recording montage. The non-inverting Sanibel TM electrode was placed on/close to the tympanic membrane of the test ear with Lectron II conductivity gel, and the inverting electrode on the contralateral mastoid with ground at F_z. A foam ear tip was placed in the ear canal to deliver stimuli and hold the TM electrode in place. Because typical tymptrode placement on or near the tympanic membrane results in extremely high impedance (for example, Henderson, 2012[37] notes that impedances > 40 kΩ are not uncommon in clinical practice) values of < 50 kΩ were considered

desirable, although recording continued even if the TM electrode impedance value exceeded this target to minimize participant discomfort with repeated placement attempts. The average impedance for the TM electrode was 50.0 kΩ (*SD* = 11.67) for the NH group (*N* = 18) and 42.13 kΩ (*SD* = 13.53) for the HI group (*N* = 10).

Alternating polarity clicks (7.1/sec; 100 μsec) were delivered to the test ear at 80 dB nHL. A recording window of 10.66 msec was used with a 2.71 msec pre-stimulus period. The responses were pre-amplified (×50,000) and filtered (10–1500 Hz). When noise exceeded 47.50 μV, data were rejected (artifact rejection was employed using this criterion). Two or three repetitions of 1024 trials were collected, resulting in a minimum of 2048 averages. Each sweep contained 256 digitized points. The EcochG waveforms were examined to determine the summating potential/action potential (SP/AP) complex. First, a baseline was assigned to which the SP and AP could be referenced. The next two adjacent positive-going peaks were identified as the SP and AP respectively. The ratio (percentage) between these two amplitudes was calculated. In instances where no clear SP or AP could be identified even after these repeated attempts, the recording was assigned a 'non-response'.

Analyses

One goal of this work was to understand which test(s) are most valuable in distinguishing presbycusis subtypes, with the premise that those tests would account for the most variance in the sample response of all the tests. In addition, we asked whether or not the multicomponent test battery would be better at grouping presbycutic subtypes than the gold standard approach. To address these aims, we employed two analytical methodologies; Exploratory Factor Analysis (EFA) and Hierarchical Cluster Analysis (HCA). Analyses were performed using SAS (version 9.4). The reader should note these analytical approaches were undertaken in solely an exploratory capacity and the small sample size limited the number of variables that could be included in the analytical models. We were unable to evaluate every test outcome for which data were collected. In the extensive data gathered from tests such as the ABR and Békésy threshold tracking, reductions were made to minimize the number of metrics that were used in the analytical models while maintaining some model stability. Test variables in the models included: Thresholds (0.25–8 kHz in dB HL attained from the standard clinical procedure), HAF (acquired from the Békésy tracking procedure), speech outcomes (WRS and WIN), DPOAEs ($2f_1$-$f_{2(Low)}$, $2f_1$-$f_{2(High)}$, f_2-$f_{1(Low)}$, f_2-$f_{1(High)}$), ABR wave V threshold, and EcochG SP/AP ratio. Individual thresholds at extended high frequencies, fine-grained DPOAE measures at every binned frequency, and ABR latencies and amplitudes were therefore not included. The limited statistical treatment is purely exploratory and represents a modeling approach to data mining we plan to perform in future studies with much larger data sets. With these points in mind, one should note the results are not yet generalizable.

Prior to inclusion in the EFA, missing EcochG data were imputed based on all the data from the model (*N* = 2 HI participants). These two missing values were imputed using the expectation-maximization (EM) algorithm implemented in SAS (proc mi). For case 1, this resulted in an imputed SP/AP ratio of 0.44 and for case 4, an imputed ratio of 0.43. The imputed values were considered only for EFA (they were not used in HCA).

Nine EFAs with oblique rotation were performed to explore the variance in participant responses attributable to each test. EFA allows one to analyze the total variance among all the unique test results considered and is a method of reducing the data into related sub-volumes. The number of factors was based on the scree plot with eigenvalues greater than 1.0 considered. Factor analysis with principal axis factor as the method of extraction and an oblique varimax rotation using squared multiple correlation for the diagonal of the correlation matrix was used to examine relationships among several test batteries and ultimately reduce the number of test variables. One goal in using this method was to examine the complete test battery, which consisted of all the tests, and reduce these extensive measures to a smaller number of relevant factors.

Second, a hierarchical cluster analysis (HCA) was carried out to find homogeneous clusters of cases based on measured characteristics. In this process, each case begins as one cluster and distances between clusters are calculated using the average linkage method; single-case clusters are turned into two-case clusters to replace the old single-case clusters. This iterative process continues until all observations are grouped into a single large cluster. The graphic result of HCA is dendrograms (or trees) that represent clusters of similar groups of participants. Because HCA was used in an inquisitive manner, traditional quantitative metrics of distance between data clusters in the dendrogram (e.g., branch lengths) were not considered here and dendrograms were examined from a more qualitative perspective. Note that both EFA and HCA were carried out on the complete data set; that is, data from NH and HI participants were not separated prior to analytical treatment.

Results

General findings

Figure 1 illustrates behavioral thresholds in dB HL (upper) and dB SPL (lower). Individuals with hearing loss (colored lines) are contrasted with the NH group (mean, black line with 95% confidence interval, grey). The upper panels represent clinical thresholds up to 8 kHz. Missing data points for thresholds in dB SPL in the lower panels represent frequencies at which no responses were obtained even at the limits of our equipment. Therefore the last available data point marks the HAF for that individual. The HAF was noticeably lower than 20 kHz in all the individuals with HI, with the exception of participant 10.

Average DPOAE amplitude as a function of f_2 (black line; $2f_1$-f_2, left and f_2-f_1, right), along with 95% confidence intervals (gray shaded zone), are shown for the NH group in Figure 2. In addition, emissions from each HI participant are displayed in panels grouped by audiometric phenotype. For the HI cases, open symbols are used where the average DPOAE levels were not separated from the noise floor by 6 dB or more. Large $2f_1$-f_2 DPOAEs were recorded in the NH group and the HI group had reduced, but sometimes present, emissions below 10 kHz (filled symbols). At higher frequencies, emissions were primarily absent (except participant 10, NH: THIN). In contrast, f_2-f_1 DPOAEs were (1) lower in amplitude than $2f_1$-f_2 even in NH individuals and (2) absent in all participants with HI at all frequencies.

Example brainstem responses (amplitude [μV] as a function of time [msec]) are depicted in Figure 3 (representative NH exemplar [23 year-old female], left and participant 7, right). Waveforms are shown for the highest-level click (80 dB nHL, upper) and at threshold (lower). Wave V latency increased with decreasing stimulus level for both example participants. Waveform morphology for participant 7 is not as clear as that of the NH subject. With the exception of participant 8, all individuals with HI had ABR thresholds that exceeded (were poorer than) their PTAs by at least 5 dB [Table 2].

SP/AP ratios were derived from the electrocochleograms as described above. Example EcochG waveforms are depicted in Figure 4 [axes and participants same as Figure 3]. This figure exemplifies the clarity of the electrocochleogram in a typical NH subject (SP/AP ratio = 0.31, left). In contrast, participants with HI often had noisier recordings and elevated ratios, such as participant 7 [Figure 4, right]. Three HI participants demonstrated elevated SP/AP ratios (participant 2, SP/AP = 0.77; participant 7, SP/AP = 0.65; and participant 8, SP/AP = 0.61) and two (participants 1 and 4) had no EcochG responses.

Table 2 presents all behavioral (upper) and physiological (lower) data for NH participants (mean and standard deviation) and individuals with HI. Participants with HI (except 10) demonstrated elevated PTAs though most had WRS within the 'good' to 'excellent' range. Highest audible frequencies ranged from 16–20 kHz for the NH group (mean = 17.7 kHz, *SD* = 1.31 kHz) and 8–19 kHz for the participant group [see Table 2 for individual results]. Participant 8 had an especially low HAF (8 kHz). The WIN 'thresholds' are considered normal if they are between –2 and 6 dB SNR. Participant 10 demonstrated a near-normal threshold even though

Table 2: Summary of behavioral (upper) and physiological (lower) test findings from the NH and HI participants. DPOAE SNR is given for each DP type for low (≤8 kHz) or high (>8 kHz) f_2 frequencies

Test	NH mean	NH SD	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8	Case 9	Case 10
Behavioral tests												
PTA (dB HL)	5.02	3.45	20.0	20.0	18.33	16.67	18.33	16.67	11.67	18.33	30.0	6.67
SRT (dB SPL)	20.79	3.72	35.0	45.0	40.0	40.0	40.0	35.0	30.0	30.0	50.0	25.0
WRS (%)	98.95	2.19	92.0	96.0	72.0	96.0	84.0	96.0	100.0	100.0	96.0	92.0
HAF (kHz)	17.7	1.31	12.5	11.2	12.5	14.0	12.5	12.5	12.5	8.0	12.5	19.0
WIN (SNR)	7.06	3.68	11.6	8.4	10.8	6.0	7.6	6.8	6.8	10.0	5.2	6.0
Physiological assays												
$2f_1$-$f_{2\,(low)}$	22.83	7.32	9.94	9.01	7.68	7.07	7.69	11.84	13.20	13.86	6.49	11.87
$2f_1$-$f_{2\,(high)}$	20.59	7.16	0.19	2.51	−0.33	1.20	2.91	5.10	2.22	0.93	2.08	20.09
f_2-$f_{1\,(low)}$	3.66	3.83	1.85	1.20	−1.53	−0.40	−1.18	0.18	−1.15	−0.03	−1.57	−0.52
f_2-$f_{1\,(high)}$	7.56	4.65	1.01	3.18	1.36	3.07	0.85	1.31	4.35	0.41	6.14	3.51
ABR wave V Threshold (dB nHL)	14.47	6.67	40.0	50.0	45.0	30.0	40.0	25.0	35.0	20.0	25.0	20.0
EcochG SP/AP ratio	0.37	0.13	NR	0.77	0.38	NR	0.4	0.35	0.65	0.61	0.47	0.08

ABR: Auditory brainstem response; dB SPL: Decibels sound pressure level; EcochG: Electrocochleography; HAF: Highest audible frequency (kHz from dB SPL thresholds); NH: Normal hearing; NR: No response; PTA: Pure-tone-average (average threshold in dB HL at 0.5, 1, and 2 kHz); SP/AP: Summating potential/action potential; SRT: Speech reception threshold; WIN: Words in Noise Test; WRS: Word recognition score

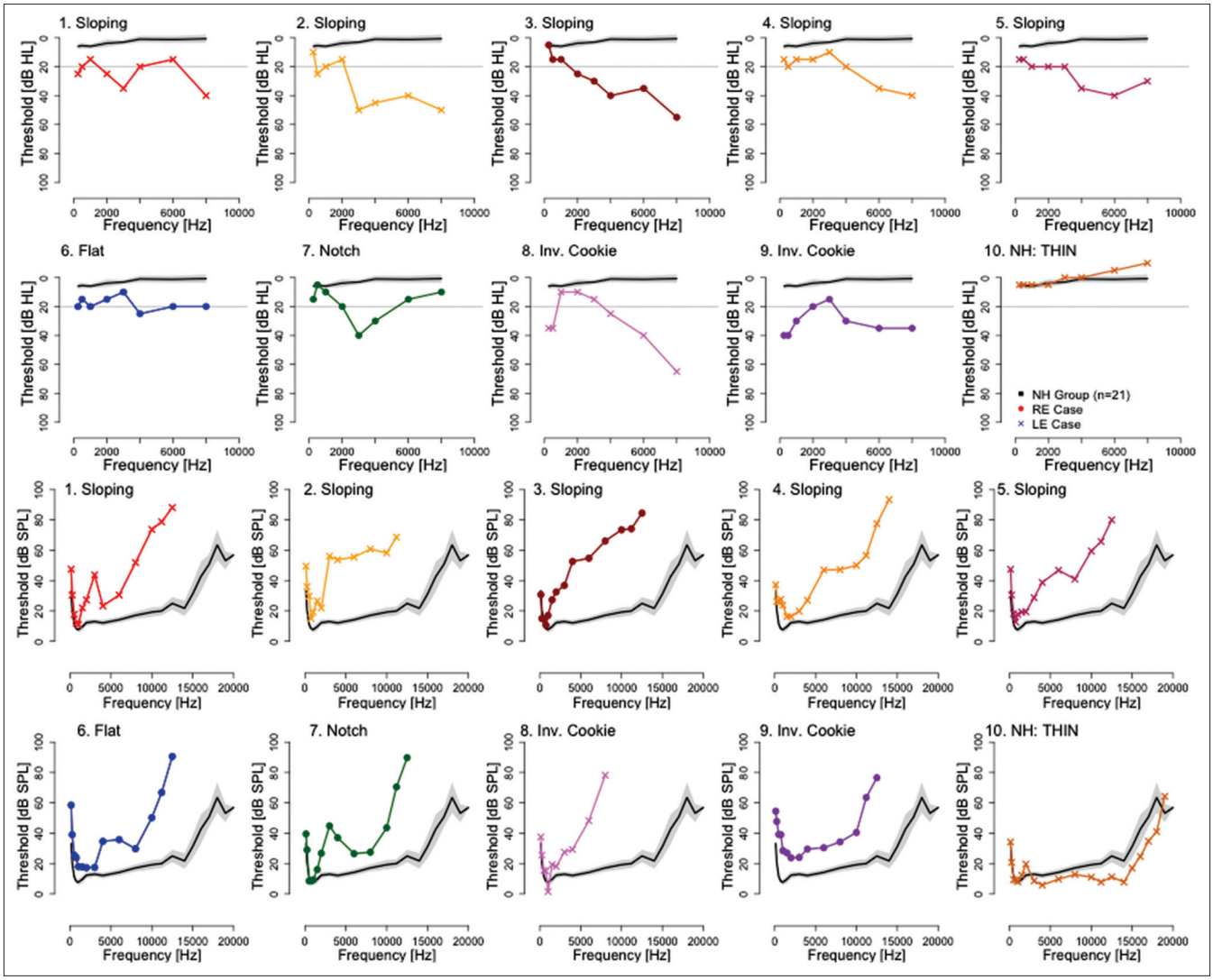

Figure 1: Behavioral thresholds in dB HL (upper panels) or dB SPL (lower panels) plotted as a function of frequency (Hz). Mean thresholds (black lines) and 95% confidence intervals (grey shaded) are also shown for the NH group (*N* = 21 for dB HL and 20 for dB SPL). Horizontal dark grey line at 20 dB HL indicates the lower bound of normal hearing. Thresholds for HI participants presented by test ear (right, circle and left, 'X'). The audiometric phenotype (based on dB HL thresholds) for each of these individuals is indicated

she complained of significant hearing difficulties in noise. Many participants had WIN scores within ± 1 SD of the NH mean (7.06 dB SNR) and only two (participants 1 and 3) had WIN thresholds falling in the 'moderate' SNR loss category.

Table 2 (lower) also shows physiological results, including DPOAEs, ABR wave V threshold, and SP/AP ratio. For DPOAEs, the value shown is in dB SNR. In NH participants, the mean $2f_1$-f_2 SNR for both the low (f_2 ≤8 kHz) and high (f_2 >8 kHz) frequency regions was above 20 dB, with an average high-frequency SNR of 22.83 dB and low-frequency SNR of 20.59 dB. In contrast, all participants with HI had greater reduction in $2f_1$-$f_{2(High)}$ than $2f_1$-$f_{2(Low)}$ (with the exception of participant 10). Overall, $2f_1$-f_2 amplitudes were lower in participants with HI compared to those with NH, with more reduction occurring at high frequencies. The SNR of f_2-f_1 was much lower than the SNR of $2f_1$-f_2, even for NH individuals. The DPOAEs from participants with HI had very poor SNRs, with the majority being <3 dB. The f_2-$f_{1(Low)}$ SNR for participant 9 was near that of the NH group [also see Figure 2], although this should be interpreted with caution as she had no measurable thresholds at these frequencies [Figure 1]. In summary, both emission types were reduced in participants with HI, with f_2-f_1 being essentially absent. However, the poor SNR for f_2-f_1 in all ears clouds easy interpretation of these findings.

Modeling results

Anecdotally, we can report that physiological approaches seemed to be more sensitive to hearing loss than

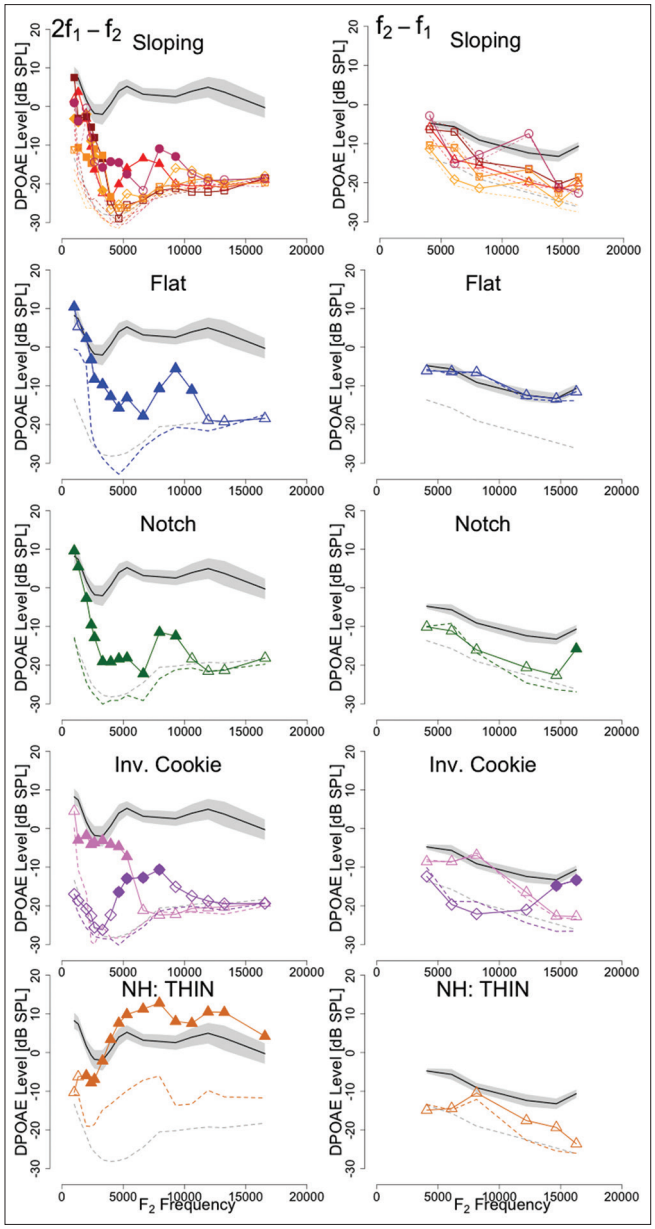

Figure 2: DPOAE levels (dB SPL) for each HI participant by DP order ($2f_1$-f_2, left and f_2-f_1, right) as a function of f_2 (Hz). Black and gray lines depict mean DPOAE and NF levels, respectively for the NH group. The shaded region represents the 95% confidence interval. Data from HI individuals are shown in various colors [see audiograms, Figure 1]. Open symbols show all binned data and closed symbols, only points with post-binning SNR of 6 dB

traditional behavioral tests (thresholds through 8 kHz and WRS). Though the standard approach could be used to identify the presence of sensory/neural HI it seemed insensitive to presbycusis classification; that is, this method did not produce distinct clusters of subjects with presumably similar pathophysiology underlying their presbycusis. Our approach permitted comparison of different test combinations to determine if any were better able to distinguish between NH and HI ears and identify subtleties in pathology that the gold standard assessment could not. It should be noted that the results presented here are investigative in nature and may not directly correlate with clinical judgments due to the limited number of patients with HI. We preliminarily addressed (1) which tests from the complete battery predict most of the variance in overall performance and (2) how a multivariate test battery could differentially cluster the 31 participants in a way the traditional battery could not.

First, nine test batteries were explored using EFA with the goal of reducing the multivariate test battery to tests most salient for distinguishing NH from pathologic ears. The details of the variables included in these models are shown in Table 3. Model 1 is the standard test battery (thresholds in dB HL from 0.25–8 kHz and WRS) and model 5, the complete battery (ABR wave V + HAF + SP/AP ratio + Thresholds [0.25–8 kHz] + WIN + WRS + $2f_1$-$f_{2(\text{Low, High})}$ + f_2-$f_{1(\text{Low, High})}$). A proxy version (model 6) was also created using PTA instead of individual thresholds from 0.25–8 kHz. This reduced the number of variables from 17 to 10 resulting in greater model stability and more reliable results. The remaining six models were composed of various combinations of tests [Table 3]. In each column, a model (a collection of tests) is presented and the factor responsible for most of the variation among responses is indicated in bold text (F1). Factors accountable for less variance are indicated as F2, F3, etc., and represent less important contributions.

Model 1 (the "gold standard" approach) shows that thresholds at 0.25 and 0.5 kHz were the highest loading items (with primary factor loadings > 0.70). Mid-frequency thresholds and WRS were in the lowest factor (F3). Model 5 was comprehensive, containing every test (except PTA). It produced seven factors, with $2f_1$-$f_{2(\text{Low})}$ and f_2-$f_{1(\text{Low, High})}$ in F1. Because the number of variables was high (17) and the number of participants low (N = 31), we examine the reduced model (6) here. It produced four factors, the principal one containing ABR wave V threshold, HAF, PTA, and $2f_1$-$f_{2(\text{High})}$. This first factor accounted for 70.2% of the overall variance in the response. This is consistent with model 1, which revealed the importance of low frequency audiometric thresholds (0.5 kHz) and models 3 and 4 which both included ABR wave V and HAF in F1. Models 3, 4, and 6 placed $2f_1$-$f_{2(\text{High})}$ in F1 as well. Last, we created model 9, a battery containing tests that seemed accountable for most of the variance (model 9 was a collection of tests which were most likely to appear in F1 in the other models). Model 9 was similar to 5 (the complete test battery) with the exclusion of WIN and SP/AP ratio. These model results were consistent with the

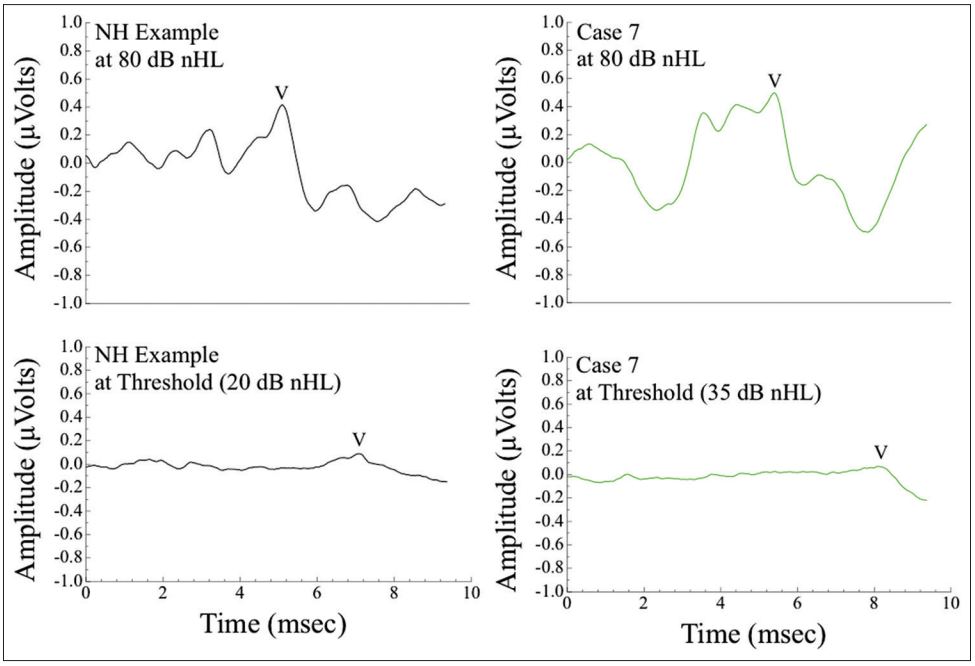

Figure 3: Example ABR waveforms (amplitude [μV] versus time [msec]) for one NH (left) and one HI (7, right) participant. Responses shown for 80 dB nHL (upper) and threshold (lower). Average of at least two repeatable runs

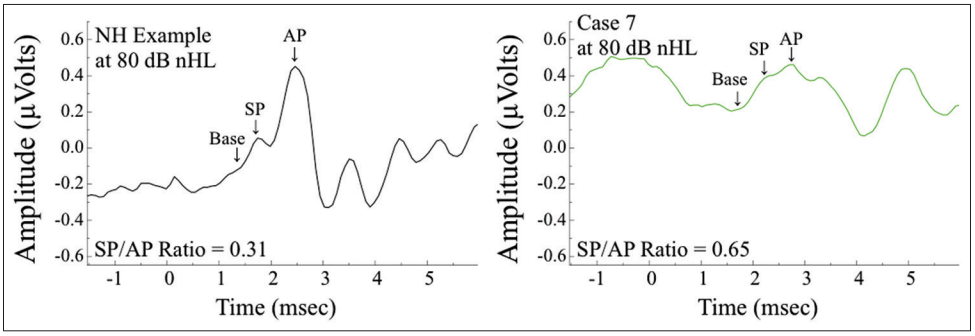

Figure 4: Example EcochG waveforms evoked by 80 dB nHL clicks. Axes and participants same as Figure 3. The NH participant demonstrates a normal SP/AP ratio (left); in contrast, HI participant 7 has an elevated SP/AP ratio (0.65; right)

others: HAF, PTA, and $2f_1$-$f_{2(High)}$ were important factors (F1; 85.3% of total variance).

In summary, examination of nine models using EFA revealed three distinct factors (HAF, PTA, and $2f_1$-$f_{2(High)}$) accounted for a large percentage of the variability in hearing outcomes among our 31 participants. Note that although nine models were evaluated, only three were discussed above due to space limitations. See Table 3 for details of the other models.

Models 1, 6, and 9 were investigated further using HCA. Cluster analysis produces dendrograms, or hierarchical branched structures containing clusters. The distinctiveness of each cluster is represented by the amount of space between adjacent clusters. As such, similar between- and within-distances between adjacent clusters denote data that are not naturally grouped. Earlier clustering on the dendrogram (lower on the graph) indicates a higher degree of commonality and longer branches represent greater dissimilarity between clusters. Figure 5 shows the resultant dendrogram for model 1 (thresholds 0.25–8 kHz and WRS). Each 'N' references an individual NH subject and participants with HI are denoted by their participant number (1–10) and color-coded for straightforward reference to Figure 1. HCA using the variables in model 1 (the "gold standard") produced a dendrogram with two clusters; the first (right side, labeled "NH Group") contained all NH participants and participant 10 (NH: THIN). Being that this model only included thresholds through 8 kHz and WRS, the finding that participant 10 clustered with the NH group is not surprising, as this participant had excellent hearing sensitivity and a WRS of 92%. This result provides evidence that the model appropriately

Table 3: Results of EFA for nine models. Bolded F1 indicates the factor that explains the most variance among a given set of tests. Model details below

Variable	Model 1*	Model 2	Model 3	Model 4	Model 5 full	Model 6* (proxy model 5)	Model 7	Model 8	Model 9*
PTA (dB HL)				F1		F1			F1
WRS (%)	F3				F7	F4	F3	F6	F3
Th. 250	F1				F3			F5	
Th. 500	F1				F6			F5	
Th. 1000	F3				F4			F4	
Th. 2000	F3				F4			F6	
Th. 3000	F2				F2			F1	
Th. 4000	F2				F6			F1	
Th. 6000	F2				F6			F2	
Th. 8000	F2				F6			F2	
HAF (kHz)		F1	F1	F1	F3	F1	F1	F2	F1
WIN (SNR)		F1	F2	F3	F5	F4			
$2f_1$-$f_{2\ (low)}$			F1	F2	F1	F2	F2	F3	F2
$2f_1$-$f_{2\ (high)}$			F1	F1	F3	F1	F1	F5	F1
f_2-$f_{1\ (low)}$		F1		F2	F1	F2	F2	F3	F2
f_2-$f_{1\ (high)}$		F1		F2	F1	F2	F2	F3	F2
ABR wave V Threshold (dB nHL)		F1	F1	F1	F2	F1	F3	F1	F3
EcochG SP/CAP ratio			F2	F3	F7	F3			

ABR: Auditory brainstem response; EcochG: Electrocochleography; EFA: Exploratory factor analysis; HAF: Highest audible frequency (kHz from dB SPL thresholds); NH: Normal hearing; NR: No response; PTA: Pure-tone-average (average threshold in dB HL at 0.5, 1, and 2 kHz); SNR: Signal-to-noise ratio; SP/CAP: Summating potential/compound action potential ratio; SRT: Speech reception threshold; Th.250: Threshold in dB HL at 250 Hz (same for other frequencies); WIN: Words in Noise Test; WRS: Word recognition score. *Model 6 is a reduced version of model 5 (thresholds at individual frequencies replaced by one value, the PTA)

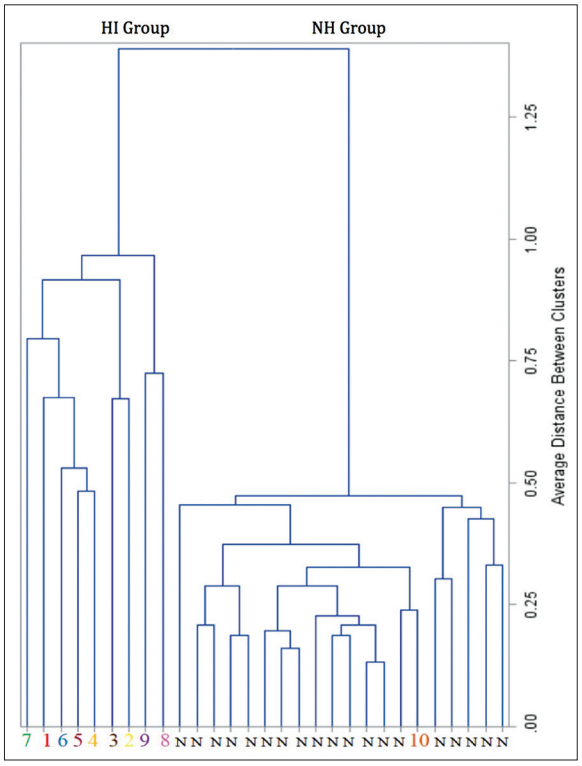

Figure 5: Dendrogram, model 1 (thresholds 0.25–8 kHz and WRS; *N* = 29). 'N' denotes NH individuals; participants with HI are presented by number [color coded to match Figure 1]. Two primary clusters are observable – NH and HI (except 10, the NH: THIN participant)

identified commonalities among participants. Second, a cluster comprised of the participants with HI was generated (*N* = 9; left side, labeled "HI Group"). This limb was composed of many smaller branches that led almost individually to each participant with HI. Interesting was the finding that participants with HI became highly separated; they were not readily parsed into homogenous groups. Using a between-cluster distance criterion of 1.0 (arbitrary selection to define distinctiveness) only two clusters could be separated: HI and NH. That is, model 1 identified participants with clinical hearing loss (behavioral thresholds ≥20 dB HL at any frequency). Such is akin to what might occur clinically in that participants with HI would be diagnosed with 'sensorineural hearing loss' but would not be segregated according to presbycusis classification.

HCA was also performed for models 6 and 9, which resulted in the same dendrogram structure. Thus, we only show the dendrogram for model 9 here. Examination of the model 9 dendrogram [Figure 6] reveals three noteworthy points. First, this battery was also able to distinguish NH (right branch) and HI (left branch) participants, but it contained far fewer branches, a finding that is especially apparent in

the HI group. Among those with HI, the dendrogram branched into two main clusters (participants 4, 3, 1, 5, 6, 7, 9, 2 and 8), with a between-cluster distance >1.0. Last, the dendrogram showed some sub-grouping among the NH participants but unlike model 1, the branches did not lead to individual cases. That is, Figure 6 shows the multivariate test battery is able to find commonalities, even among NH participants, that distinguish them from others. A final point is worthy of mention. Model 9 highlights the distinctiveness of three participants: 8 (with between-clusters distance of >1.0) and 2 and 4 (between-clusters distance >0.25). Observation of the dendrogram reveals longer branches leading to these three individuals, suggesting they are somewhat dissimilar from the other hearing impaired individuals. In light of this dendrogram result, potential differences between these three participants and the other HI individuals will be discussed below.

Discussion

A method using behavioral hearing thresholds through 8 kHz and WRS is commonly employed to diagnose presbycusis. However, this clinical "gold standard" method does not allow the differential categorization of presbycusis due to lesions of the stria vascularis, hair cells, or spiral ganglion. Neither does this method allow a finer-grained differentiation between presbycusis caused purely by aging and that due to a variety of other factors (e.g., genetics, ototoxic drugs, cardiovascular disease, and noise). Attempts at differential diagnosis, sometimes post-mortem, have been provided by a number of studies[4,6,8,13,15] and although isolated case reports have been able to link audiometric shapes from living patients with histopathological findings,[4,6,38] human temporal bone studies have, in general, suggested that the audiogram is inadequate for discerning site(s) of lesion.[16,17] With prophylactic or therapeutic strategies specific to different sites of lesion on the horizon,[21,22,26,27] enhanced diagnostic methods identifying specific etiologies and sites of lesion are needed. Here, we explore our statistical findings and the unique features of presbycutic ears they identified. We present a preliminary framework for improving upon the accepted method of diagnosing this highly prevalent disease.

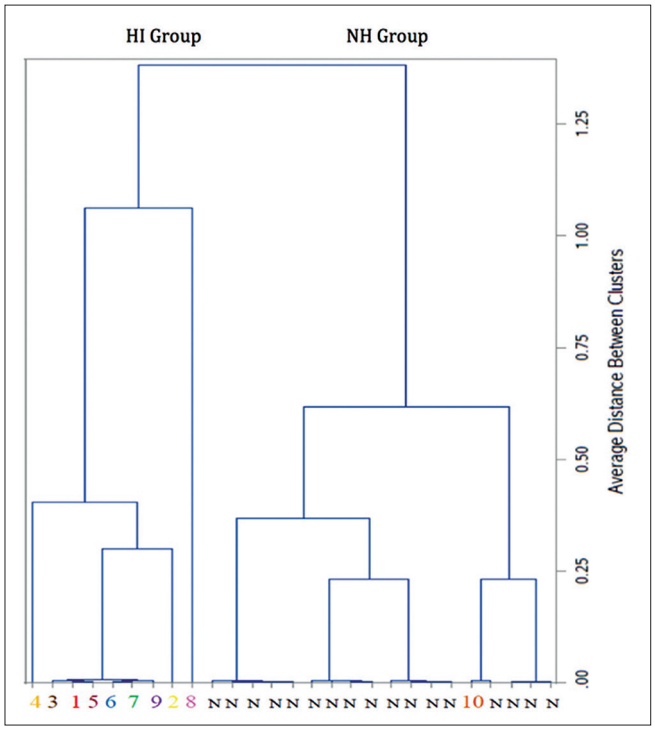

Figure 6: Dendrogram, model 9 (same as result for model 6). Two clusters are seen, as in Figure 5, but in the HI group, participants 8, 2, and 4 are distinct. Interestingly, HI participants do not cluster in a way the audiometric phenotypes would predict (e. g., the red-orange numbers do not cluster together)

The results of this work suggest some measures of auditory function (namely HAF, PTA, and $2f_1$-$f_{2(High)}$) may be more advantageous for differential categorization of presbycutic pathology than others. These tests were determined using EFA [Table 3; see F1 variables] and are especially promising for widespread application because they are available in most audiology clinics. The number of tests considered was, to our knowledge, the most extensive of any comparable study to date. We purposefully included objective measures of cochlear (DPOAE and EcochG) and neural (ABR) auditory function, as well as a number of behavioral metrics (extended high frequency audiometry and speech testing). In spite of sample size limitations, we have shown that our comprehensive test battery might be better at segregating presbycutic ears into unique categories compared to the traditional method. The statistical approach undertaken here suggests some tests are potentially more valuable than others in this regard. Despite the extensiveness of our test battery, it was certainly not all-inclusive and should not be regarded as a definitive approach to phenotyping moving forward. Rather, this work serves as a pilot study and a starting place from which the removal of some tests and addition of others might be most suitable. For example, results suggest EcochG might not be useful in differentiation of peripheral sites of lesion in presbycusis, but ABR might be valuable. A more comprehensive approach would incorporate additional aspects of the ABR (e.g., waves I and V latency, amplitude, and thresholds) and possibly, middle or long latency potentials to rule

out central pathology. In addition, because many of the participants in this study had sloping audiometric configurations, examination of behavioral pure-tone averages calculated using other frequencies might be helpful (e.g., average of 1, 2, and 4 kHz or 2, 4, and 8 kHz). Inclusion of these variables might add valuable information to the differential categorization scheme attempted herein and are planned for future studies with larger numbers of participants.

The identification of key tests of auditory function (HAF, PTA, and $2f_1$-$f_{2(High)}$) is a valuable first step in understanding presbycusis subtypes and the optimal way to identify them. The results of the EFA lead to the use of HCA to understand the categorization of individuals using particular test combinations. Cluster analysis revealed that although the gold standard approach and our multivariate test batteries could both distinguish NH from HI, the enhanced routine might be superior because it identified homogeneities among participants without needless individualization [Figures 5 and 6]. The HI participants in this study had various combinations of reduced DPOAEs, elevated ABR wave V thresholds, elevated SP/AP ratios, and non-optimal WRS. The PTAs were often normal or near normal [Table 1]. Results of the HCA suggest that among the individuals with HI, one was particularly distinctive (participant 8) and two (2 and 4) stood out to a lesser extent.

Here, we attempt to elucidate the site of lesion(s) in these individuals (participants 8, 2, and 4) in light of the statistical findings. The reader should note that the small sample size necessitates cautious interpretation of our understanding of these data. Further, we stress that in the absence of corroborative histopathology, our scheme and resultant interpretation is not a definitive means of differential categorization. Rather, we attempt to use the results of the HCA to explore behavioral and/or physiological differences between clusters. At this point it is a qualitative strategy and one that will be greatly enhanced with a larger sample in which the findings from such clusterings may be more illuminating.

We first consider participant 8, a 70-year-old female with occasional tinnitus, history of smoking, and hypertension. She had the most severe hearing loss of all the participants, with a HAF of 8 kHz. Her residual mid-frequency hearing likely permitted excellent performance on word recognition testing (100% WRS) and only a 'mild' SNR loss on the WIN test. Her physiological measures are quite interesting and their exclusion would yield an incomplete representation of her auditory profile. The click-evoked ABR wave V threshold of 20 dB nHL, in combination with the behavioral speech results, suggest minimal neural degeneration, at least up to the inferior colliculus. In contrast, the EcochG showed an elevated SP/AP ratio (0.61), and DPOAEs (both types) were absent above f_2 of 6 kHz. Taken together, these findings point toward a form of cochlear hearing loss, *possibly* of the strial subtype. The participant also reported hypertension, which may be relevant to the presbycusis classification, as hypertension is a risk factor for cardiovascular disease that has also been associated with hearing impairment.[39-41] Hypertension could conceivably reduce the endocochlear potential via a reduction in blood flow to the stria vascularis, thereby providing some support for cochlear presbycusis of the strial subtype. One caveat to this preliminary conclusion is the severity of the loss (>60 dB HL at 8 kHz); it has been suggested that cochlear losses of metabolic origin can not exceed approximately 60 dB HL and greater losses are due to the combined effects of reduced endocochlear potential and hair cell degeneration.[15] Though this participant did not report noise exposure, OHC loss due to aging remains a possibility, especially as the loss occurred in the high frequencies (coincident with basal hair cell loss associated with aging as seen in gerbils).[42] Though participant 8 did not report vertigo or aural fullness, her history of tinnitus and audiological profile (especially the elevated SP/AP ratio) hint at the possibility of concomitant strial disease, confounding a straightforward determination of site of lesion.

The HCA also highlighted the exceptionality of participants 2 and 4. Both had 'sloping' sensory/neural hearing losses. Participant 2 was a 64-year-old male with self-reported noise exposure, tinnitus, and vertigo. Performance on some of the tests from the multivariate test battery was distinct from the other HI individuals. For example, he had the highest (worst) ABR threshold of all the participants (50 dB nHL). The ABR reflects spiral ganglion population and neural synchrony, both of which may decline with increased age.[36,43] Being that the PTA was within normal limits (20 dB HL), the elevated ABR threshold points toward a primary classification of neural presbycusis (spiral ganglion lesion). However, it is important to note the likelihood that factors in addition to aging (namely, noise exposure) contributed to the development of this hearing loss. Noise exposure has been shown to aggravate ARHL in animal models[44] and if the exposure was long-term, hair cell damage would be expected.[45] Further justification of a mixed presbycutic classification for this participant can be found upon examination of the DPOAE responses. The high-frequency SNR of both DPOAEs was ≤3.5 dB although $2f_1$-$f_{2(Low)}$ had an SNR of ~9 dB

[Table 2, Figure 2]. Taken together, the auditory profile suggests a combined sensory/neural categorization, perhaps with a dominant neural component. If true, it is interesting that the WRS was minimally affected (at 92%), but this may reflect the ability of persons with near-normal PTAs to perform well on this task because speech information is conveyed primarily at lower frequencies and multiple redundancies are available in the speech material.

Participant 4, a 56 year-old male, was also partially segregated from the HI branch [Figure 6]. He reported a heart attack, but otherwise an unremarkable case history. This information, although not included in the analytical models, does provide a hint as to the possible site(s) of lesion. The ABR threshold was elevated above the NH group [30 dB nHL, see Table 2] but similar to many of the others from the HI group. The $2f_1$-f_2 DPOAEs amplitudes were some of the lowest (poorest) among the hearing impaired participants, although the f_2-f_1 SNR was comparable to the others. Interestingly, the HAF (14 kHz) was one of the highest (best) of all the HI participants (excluding 10, NH: THIN). An initial conclusion might be that this participant was separated by HCA because the pathology underlying his presbycusis may be less severe or at an earlier stage. Given the available data, we conjecture participant 4 likely presented with a mixed pathology, perhaps dominated by cochlear contributions. More specifically, the DPOAE responses point to outer hair cell damage. Of course, the interpretation of these cases is preliminary given the small number of subjects in the HI group. The real strength of statistical approaches such as HCA could be harnessed by considering many cases and determining which auditory tests are most valuable for differential categorization. The present work and our understanding of the findings is essentially a pilot attempt to explore the feasibility of a larger-scale approach.

This work is not without limitations, including the small sample. We evaluated an exhaustive test battery consisting of behavioral and physiological indices of auditory function. From a qualitative standpoint, these data were sufficient to arrive at a number of interesting conclusions. But from a more analytical perspective, additional data are needed to carry out factor and cluster analyses with greater confidence. Second, though we used measures of both cochlear and neural function, the crux of the test battery resided in the DPOAE measurements. Great energy was invested in determining optimal stimulus parameters for the f_2-f_1 DPOAE recordings and this report is the first, to our knowledge, to present f_2-f_1 DPOAEs from damaged human ears. In contrast to the DPOAEs, the electrophysiological measures were relatively rudimentary. Future studies might evaluate the speech ABR and the role of stimulus rate on presbycusis phenotyping schemes. It is possible that such measures might be capable of revealing more about neural presbycusis than has been shown here. We found EcochG to be a relatively unimportant factor in our multivariate test battery. At present, our results do not support the use of EcochG or f_2-f_1 DPOAEs for routine presbycusis phenotyping unless additional strial dysfunction is expected (e.g., hydrops). Despite the extensive array of tests considered in the present report, any non-invasive approach for improving understanding of underlying pathology is complicated and somewhat limited. Nonetheless, this work is promising as we are able to go beyond the gold standard (audiometry and WRS) and improve our appreciation of presbycusis subtypes using behavioral and physiological approaches to assessing auditory function.

A less significant yet important issue related to participant recruitment is the male-female balance in this study. In the NH group, the majority of participants were female (16 of 21) but in the HI group, only 3 of 10 were female [Table 1]. Studies, both cross-sectional[46,47] and longitudinal,[48] have indicated differences in the prevalence and incidence of ARHL between and men and women (see Gordan-Salant, 2005[49] for review). However, whether or not such gender differences exist when ear canal acoustics are carefully considered and compensatory calibration strategies are employed has been called into question.[3] The study by Lee *et al*., (2012) presented behavioral hearing thresholds up to 20 kHz in 10–65 year olds and observed no statistically significant differences between thresholds of men and women. Our behavioral Békésy threshold tracking procedure, instrumentation, and calibration were the same as in the Lee *et al*., (2012) report. Further, our study was not an attempt to establish prevalence estimates of specific presbycusis subtypes. Nonetheless, future attempts at differential categorization of presbycusis might consider gender as a variable in statistical models or strive for a more equal distribution of men and women in the study population.

Last, an obvious limitation of this study is the inclusion of noise-exposed individuals in the HI group. Hearing loss due exclusively to aging is difficult to study in industrialized societies due to the ubiquity of noise exposure. Presbycusis itself can be thought of as a combination of intrinsic aging processes and extrinsic auditory assaults such as noise exposure and ototoxic agents.[50] In the present report, we define presbycusis in this way and as such, individuals with noise exposure (e.g., participants 1, 2, 5, and 7) were

included for study. We opted to include individuals in this study who presented with multiple risk factors for hearing loss in addition to aging (e.g., noise exposure, hypertension, smoking, etc). This choice was motivated in part by our desire to consider a sample as close to the clinical reality of the presbycutic population as possible, as we ultimately aim to develop an approach employable in the clinic. We are thus unable to exclude noise exposure as the primary antecedent of auditory damage in some cases (e.g., participant 7, notched audiogram). It is possible that noise-exposed ears age differently than unexposed ears,[50] but from a statistical standpoint, this should have no bearing on the results. A differential classification scheme such as the one preliminarily proposed here might be useful in separating cases with noise exposure from those with 'pure' presbycusis and may clarify behavioral and physiological phenotypes with overlapping pathophysiologies. In the present report, we do not claim our statistical approach is able to distinguish a case of noise-induced hearing loss from pure presbycusis. However, with a larger data set, we plan to include noise exposure status in the statistical models as well as other self-reported data, which might permit such differentiations.

An advantage of this study is the inclusion of behavioral thresholds up to 20 kHz. We showed that slopes (configurations) observed below 8 kHz cannot necessarily be extrapolated to higher frequencies [Figure 1]. In our sampling of HI participants, hearing thresholds and DPOAEs at high frequencies (>8 kHz) were almost universally affected, consistent with the view of the cochlear base being primarily or initially affected by presbycusis. The number of participants in this report limited our statistical evaluation of extended high frequency audiometry to the HAF, but we showed that HAF is a valuable marker for identifying distinctiveness among those with HI. In the future, the slope of high frequency audiograms, as well as the corner frequency, might prove to be important markers of auditory health or indicators of specific presbycusis types. Ultimately, a reduced version of the multivariate test battery should be built. An extensive data set will be required to determine how to best tailor test selection for a given patient, perhaps using a modeling approach to resolve the appropriate collection of tests based on the audiometric phenotype or other key factors from the initial patient examination.

In summary, we set out to identify a test battery capable of differential categorization of presbycusis. The findings suggest an enhanced test battery might be better for recognition of homogeneities among participants with HI than the gold standard approach. Further, it seems to provide more precise diagnostic information that can aid in distinguishing site(s) of lesion. The study is preliminary, but the results are powerful, as they underscore the notion that routine clinical evaluation is insufficient in distinguishing pathology and therefore, will be a limiting factor in the application of emerging therapeutic approaches. The gold standard method has a longstanding tradition of use in the clinic, especially for the fitting of amplification devices, but needs to be advanced for patients to benefit from medical interventions that are site-specific. Ultimately, our multivariate test battery needs to be reduced to fit into a standard one-hour clinical appointment and verified in a much larger population (ideally, with corresponding histopathological data). It is only once this additional information is obtained that any recommendations to alter current clinical practice can be made. The main motivation behind this work was to find a means of categorizing presbycusis subtypes that will permit eventual application of targeted biological therapeutics. We determined that our more thorough test battery is better for understanding the complex etiology of presbycusis in impaired ears, indicating that in the long-term, it might promote earlier detection of presbycusis and better appropriation of treatments. In the future, these phenotyping techniques can be applied on an epidemiological scale and ultimately, might inform suitable treatment methodologies for each presbycusis subtype. The public health implications of this work are thus potentially quite far-reaching.

Acknowledgment

We wish to express our gratitude to Dr. Diana Callesano and Ms. Emily Kahn for assistance with data collection. In addition, we thank Mr. Chun Liang Chan for his software and data management support. Additional thanks are due to Drs. Mary Ann Cheatham, Jason Sanchez, Gayla Poling, and Jonathan Siegel and Mr. James Dewey, for their input. Last, this research would not have been possible without the time and commitment of the research participants, to whom we owe much appreciation.

References

1. Kochkin S. MarkeTrak VII: Hearing loss population Tops 31 million people. Hear Rev 2005;12:16-29.
2. Arlinger S. Negative consequences of uncorrected hearing loss - A review. Int J Audiol 2003;42:2S17-20.
3. Lee J, Dhar S, Abel R, Banakis R, Grolley E, Lee J, *et al*. Behavioral hearing thresholds between 0.125 and 20 kHz using depth-compensated ear simulator calibration. Ear Hear 2012;33:315-29.
4. Pauler M, Schuknecht HF, White JA. Atrophy of the stria vascularis as a cause of sensorineural hearing loss. Laryngoscope 1988;98:754-9.

5. Saremi A, Stenfelt S. Effect of metabolic presbyacusis on cochlear responses: A simulation approach using a physiologically-based model. J Acoust Soc Am 2013;134:2833-51.
6. Schuknecht HF, Gacek MR. Cochlear pathology in presbycusis. Ann Otol Rhinol Laryngol 1993;102:1-16.
7. Makary CA, Shin J, Kujawa SG, Liberman MC, Merchant SN. Age-related primary cochlear neuronal degeneration in human temporal bones. J Assoc Res Otolaryngol 2011;12:711-7.
8. Schuknecht HF. Further observations on the pathology of presbycusis. Arch Otolaryngol 1964;80:369-82.
9. Schuknecht HF. Auditory and cytocochlear correlates of inner ear disorders. Otolaryngol Head Neck Surg 1994;110:530-8.
10. Schuknecht HF, Watanuki K, Takahashi T, Belal AA Jr, Kimura RS, Jones DD, *et al*. Atrophy of the stria vascularis, a common cause for hearing loss. Laryngoscope 1974;84:1777-821.
11. Schuknecht HF, Ishii T. Hearing loss caused by atrophy of the stria vascularis. Nihon Jibiinkoka Gakkai Kaiho 1966;69:1825-33.
12. Crowe SJ, Guild SR, Polvogt LM. Observations on the pathology of high-tone deafness. Bull Johns Hopkins Hosp 1934;54:315-79.
13. Pauler M, Schuknecht HF, Thornton AR. Correlative studies of cochlear neuronal loss with speech discrimination and pure-tone thresholds. Arch Otorhinolaryngol 1986;243:200-6.
14. Allen PD, Eddins DA. Presbycusis phenotypes form a heterogeneous continuum when ordered by degree and configuration of hearing loss. Hear Res 2010;264:10-20.
15. Dubno JR, Eckert MA, Lee FS, Matthews LJ, Schmiedt RA. Classifying human audiometric phenotypes of age-related hearing loss from animal models. J Assoc Res Otolaryngol 2013;14:687-701.
16. Nelson EG, Hinojosa R. Presbycusis: A human temporal bone study of individuals with downward sloping audiometric patterns of hearing loss and review of the literature. Laryngoscope 2006;116:1-12.
17. Nelson EG, Hinojosa R. Presbycusis: A human temporal bone study of individuals with flat audiometric patterns of hearing loss using a new method to quantify stria vascularis volume. Laryngoscope 2003;113:1672-86.
18. Suga F, Lindsay JR. Histopathological observations of presbycusis. Ann Otol Rhinol Laryngol 1976;85:169-84.
19. Mizutari K, Fujioka M, Hosoya M, Bramhall N, Okano HJ, Okano H, *et al*. Notch inhibition induces cochlear hair cell regeneration and recovery of hearing after acoustic trauma. Neuron 2013;77:58-69.
20. Lustig LR, Akil O. Cochlear gene therapy. Curr Opin Neurol 2012;25:57-60.
21. Atar O, Avraham KB. Therapeutics of hearing loss: Expectations vs reality. Drug Discov Today 2005;10:1323-30.
22. Lei D, Gao X, Perez P, Ohlemiller KK, Chen CC, Campbell KP, *et al*. Anti-epileptic drugs delay age-related loss of spiral ganglion neurons via T-type calcium channel. Hear Res 2011;278:106-12.
23. Tieu C, Campbell KC. Current pharmacologic otoprotective agents in or approaching clinical trials: How they elucidate mechanisms of noise-induced hearing loss. Otolaryngol 2013;3:1.
24. Bielefeld EC, Tanaka C, Chen GD, Henderson D. Age-related hearing loss: Is it a preventable condition? Hear Res 2010;264:98-107.
25. Izumikawa M, Minoda R, Kawamoto K, Abrashkin KA, Swiderski DL, Dolan DF, *et al*. Auditory hair cell replacement and hearing improvement by Atoh1 gene therapy in deaf mammals. Nat Med 2005;11:271-6.
26. Parker MA. Biotechnology in the treatment of sensorineural hearing loss: Foundations and future of hair cell regeneration. J Speech Lang Hear Res 2011;54:1709-31.
27. Shibata SB, Budenz CL, Bowling SA, Pfingst BE, Raphael Y. Nerve maintenance and regeneration in the damaged cochlea. Hear Res 2011;281:56-64.
28. Poling GL, Siegel JH, Lee J, Lee J, Dhar S. Characteristics of the $2f_1$-f_2 distortion product otoacoustic emission in a normal hearing population. J Acoust Soc Am 2014;135:287-99.
29. Carhart R, Jerger JF. Preferred method for clinical determination of pure-tone thresholds. J Speech Hear Disord 1959;24:330-45.
30. Roup CM, Wiley TL, Safady SH, Stoppenbach DT. Tympanometric screening norms for adults. Am J Audiol 1998;7:55-60.
31. Levitt H. Transformed up-down methods in psychoacoustics. J Acoust Soc Am 1971;49:467.
32. Wilson RH. Development of a speech-in-multitalker-babble paradigm to assess word-recognition performance. J Am Acad Audiol 2003;14:453-70.
33. Wilson RH, Strouse A. Northwestern University Auditory Test No. 6 in multitalker babble: A preliminary report. J Rehabil Res Dev 2002;39:105-13.
34. Dhar S, Talmadge CL, Long GR, Tubis A. Multiple internal reflections in the cochlea and their effect on DPOAE fine structure. J Acoust Soc Am 2002;112:2882-97.
35. Long GR, Talmadge CL, Lee J. Measuring distortion product otoacoustic emissions using continuously sweeping primaries. J Acoust Soc Am 2008;124:1613-26.
36. Boettcher FA. Presbyacusis and the auditory brainstem response. J Speech Lang Hear Res 2002;45:1249-61.
37. Henderson CC. Electrocochleography: Clinical applications. In: Atcherson SR, Stoody TM, editors. Auditory Electrophysiology: A Clinical Guide (Google eBook). Ch. 12. New York: Thieme Medical Publishers, Inc; 2012. p. 175-86.
38. Castaldo A, Linthicum FH Jr. Stria vascularis hearing loss. Otol Neurotol 2006;27:285-6.
39. Brant LJ, Gordon-Salant S, Pearson JD, Klein LL, Morrell CH, Metter EJ, *et al*. Risk factors related to age-associated hearing loss in the speech frequencies. J Am Acad Audiol 1996;7:152-60.
40. Gates GA, Cobb JL, D'Agostino RB, Wolf PA. The relation of hearing in the elderly to the presence of cardiovascular disease and cardiovascular risk factors. Arch Otolaryngol Head Neck Surg 1993;119:156-61.
41. Agarwal S, Mishra A, Jagade M, Kasbekar V, Nagle SK. Effects of hypertension on hearing. Indian J Otolaryngol Head Neck Surg 2013;65:614-8.
42. Tarnowski BI, Schmiedt RA, Hellstrom LI, Lee FS, Adams JC. Age-related changes in cochleas of mongolian gerbils. Hear Res 1991;54:123-34.
43. Konrad-Martin D, Dille MF, McMillan G, Griest S, McDermott D, Fausti SA, *et al*. Age-related changes in the auditory brainstem response. J Am Acad Audiol 2012;23:18-35.
44. Kujawa SG, Liberman MC. Acceleration of age-related hearing loss by early noise exposure: Evidence of a misspent youth. J Neurosci 2006;26:2115-23.
45. Chen GD, Fechter LD. The relationship between noise-induced hearing loss and hair cell loss in rats. Hear Res 2003;177:81-90.
46. Moscicki EK, Elkins EF, Baum HM, McNamara PM. Hearing loss in the elderly: An epidemiologic study of the Framingham Heart Study Cohort. Ear Hear 1985;6:184-90.
47. Cruickshanks KJ, Wiley TL, Tweed TS, Klein BE, Klein R, Mares-Perlman JA, *et al*. Prevalence of hearing loss in older adults in Beaver Dam, Wisconsin. The Epidemiology of Hearing Loss Study. Am J Epidemiol 1998;148:879-86.
48. Pearson JD, Morrell CH, Gordon-Salant S, Brant LJ, Metter EJ, Klein LL, *et al*. Gender differences in a longitudinal study of age-associated hearing loss. J Acoust Soc Am 1995;97:1196-205.
49. Gordon-Salant S. Hearing loss and aging: New research findings and clinical implications. J Rehabil Res Dev 2005;42:9-24.
50. Gates GA, Schmid P, Kujawa SG, Nam B, D'Agostino R. Longitudinal threshold changes in older men with audiometric notches. Hear Res 2000;141:220-8.

How to cite this article: Baiduc RR, Lee J, Dhar S. Toward a test battery for differential categorization of age-related hearing loss. J Indian Speech Language Hearing Assoc 2014;28:25-40.

Source of Support: This work was partially funded by a grant (DC008420) from the National Institute for Deafness and other Communicative Disorders of the National Institutes of Health, USA. The Northwestern University Audiology Clinic (Dr. Nielson, Clinic Director) supported the project by lending necessary equipment. **Conflict of Interest:** None declared.