hi HealthInnovations Methods for Prescribing Hearing Device Gain: Reliability and Accuracy

Dianne J. Van Tasell, Ph.D.

Dianne Van Tasell is Adjunct Professor in the Department of Speech, Language, and Hearing Sciences at the University of Minnesota, Minneapolis, MN. She is an independent consultant in hearing science. She was the lead technical consultant to IntriCon Corporation on development of the hi HealthInnovations hearing assessment methods.

Contact information:

5830 N. Placita Amanecer
Tucson, AZ  85718
520-615-5148
diannevantasell@gmail.com

Abstract:

Two versions are described of a self-test of hearing. The hi HealthInnovations Clinic Version is designed to allow self-measurement of pure tone air conduction thresholds at selected frequencies from 500 – 6000 Hz. The Home Version can be accessed via the internet, and combines data from a number of sources (questionnaire, uncalibrated measurement of selected AC thresholds, age, and gender) to estimate hearing status based on publicly-available US demographic data. Both versions are designed to identify persons who could benefit from open-fit hearing devices (symmetrical, mild-moderate hearing losses), to program open-fit devices appropriately for those persons, and to make appropriate referrals for persons who are not candidates. Data are reported on: 1) reliability and accuracy of AC thresholds measured using a calibrated beta version of the Clinic Version software; and 2) accuracy of the Home Version gain prescription methods for 144 subjects with known symmetrical audiograms.
Open-fit hearing devices do not require custom ear impressions and earmolds, and therefore the same device can fit most ears, using a limited number of tubing, connector, and tip sizes. This flexibility of physical fit has the potential to increase greatly access to hearing devices for people who are open-fit candidates. Studies in the United Kingdom have already shown that open-fit devices can make hearing health care more accessible and less costly (Smith et al., 2008).

hi HealthInnovations makes available technologically advanced open-fit hearing devices to persons who are candidates for them. Candidacy is assessed using tests that the potential hearing device user can take either in a health care provider’s office or at home via the internet, using a variety of platforms. The purpose of this article is to describe the methods developed for these self-tests and report the results of studies evaluating their reliability and accuracy.

The goals of the hiHealthInnovations hearing tests are:

1. Accurately assess the hearing of users for the purposes of identifying people who could potentially benefit from open-fit hearing devices;

2. Make appropriate audiological and medical referrals for persons who are not candidates for open-fit devices so that they may get the custom earpieces, hearing device programming, and individualized assistance that they need

3. For those users who are candidates, provide open-fit hearing devices accurately programmed for them according to methods that meet current audiological standards of practice.

**Two Versions of Hearing Test**

The **Clinic Version** will be available on iPad and in an Adobe AIR version that will run on PC and Mac computers. This version will use custom-designed earphones for which reference equivalent threshold sound pressure levels have been determined according to the methods described in ANSI S3.6 – 2010, and that may be calibrated in an HA-1 coupler. The resulting thresholds are used to program the hearing devices in one memory with the parameters prescribed by NAL-NL2 (Keidser et al., 2011). The two additional memories bracket the NAL prescription in order to accommodate individual preferences for variations from NAL.

The **Home Version** will be available on iPad, iPhone, iPod, and Android platforms as well as in a Flash version that runs on PC and Mac computers. Any commercially-available earphones may be used with the home version; insert earphones are recommended. Devices are also programmed using NAL-NL2, as described later in this article.

**Referral Criteria**
**Asymmetry.** Asymmetrical losses are detected using criteria based on differences between ears reported by Schlauch et al. (1995) when acoustic neuroma is present in one ear. If the average thresholds between 2 – 6 kHz differ across ears by >15 dB, then the user is referred to a healthcare professional for further evaluation.

**Inability to do the assessment:** During an initial practice session the user is instructed to make two threshold adjustments; if after five tries the user cannot produce two consecutive adjustments within 5 dB, he/she is instructed to either seek additional assistance, be it office personnel in the Clinic Version, or a health care professional in the Home Version.

**Loss too severe for open fit:** If average AC threshold is >40 dB HL between 500 – 1000 Hz or >65 dB HL between 2000 – 6000 Hz, then the loss is judged too severe to be fit by NAL-NL2 gain in an open-fit device (Smith et al, 2008; Van Tasell et al, 2008). The user is advised to see a healthcare professional for further assistance and evaluation.

**Measurement Method: Air Conduction Thresholds**

The development objectives for the method for measurement of air conduction (AC) thresholds were that it be:

1. Engaging and interesting for the user;
2. Reliable (test-retest reliability equivalent to that reported in the audiology literature for standard audiometry in hearing-impaired population); and
3. Accurate (measured difference between thresholds obtained by new method and standard audiometry equivalent to test-retest reliability for standard audiometry reported in the audiology literature for hearing-impaired population).

Air conduction thresholds are measured in both office and home versions of the hi HealthInnovations hearing test via a modified psychophysical method of adjustment. The user moves an on-screen adjustment tool in order to adjust the level of pure tone pulse trains as instructed by the software. In the Flash and Adobe Air versions, where the tool must be manipulated using a mouse, the adjustment tool is an “endless” slider (Figure 1), much like the horizontal volume control wheel on an old-fashioned radio. For touch-screen applications, the adjustment tool is “endless” wheel modeled after the click wheel on the original iPod (Figure 2).

Figures 1 and 2 about here

**Measurement.** After successful completion of practice the test sequence commences. At each test frequency the user is first asked if the tone sequence is audible; if not, he/she is asked to turn up the volume until it is. Then the user turns down the volume until the tones are no longer audible; at this point the software further reduces the volume by a random amount so as to remove any on-screen cues regarding the level of the tones. Then the user is asked to adjust the volume back up until the tones can just barely be heard. The process is repeated up to five times until two adjustments within 5 dB are obtained. Threshold is computed as the mean of these two. The informed reader will recognize this method as an application of the Carhart-Jerger (1959) method to the method of adjustment: threshold adjustments are always made on
ascending runs. In the Clinic Version AC thresholds (in dB HL) are measured at 500, 1000, 2000, 3000, 4000 and 6000 Hz in each ear.

**Evaluation of Air Conduction Measurement Method (Clinic Version).** Feasibility studies of the method-of-adjustment threshold measurement implemented on an iPad were carried out with a beta-test version of the Clinic Version software. Data collected from 25 hearing-impaired subjects at the University of Western Ontario and 30 hearing-impaired subjects at the University of Arizona are reported here. All studies were approved by the participating institution’s Institutional Review Board (IRB). Across locations, subject age ranged from 21 – 86 years, with a median age of 71. All subjects had known sensorineural hearing loss in at least one ear. Average PTA across individual ears (500, 1000, 2000 Hz) was 40 dB HL, with minimum of 8 dB HL and maximum of 92 dB HL. At both locations the method evaluated was the (beta-test) Clinic Version, running on an iPad and using Etymotic ER-3A insert earphones. Both the audiometer and the iPad were calibrated to ANSI S3.6-2010 specifications using an HA-2 coupler. Each subject completed several hearing test conditions including: 1) measurement of air conduction thresholds using standard clinical procedures using a calibrated clinical audiometer and ER-3A earphones; 2) two repetitions of self-test thresholds in each ear using the method of adjustment on a calibrated iPad. If there was an audiogram on file for a subject that had been taken during the previous six months, then that audiogram was used as the standard audiogram. Subjects did not receive special instruction with the iPad other than the standard user instructions and practice that appeared on the screen. Supervisors of testing were allowed to answer subjects’ questions, but not to help them make threshold adjustments. At each research location, one subject could not complete the iPad self-assessment because he/she could not make consecutive adjustments within the required 5 dB. One of these subjects was 84 years old; the other was 77.

I. **Reliability.** Test-retest reliability is quantified by the difference between the first and second thresholds measured. The usual clinical standard is difference ± 10 dB; this corresponds roughly to the range of measurement error inherent in standard audiometry, including calibration tolerances (Schmuziger et al., 2004). Ninety-nine percent of subjects’ second iPad thresholds were within 10 dB of the first; this compares favorably even with test-retest data from young otologically normal adults, shown for comparison in Table 1, and with test-retest data from one study in which the subjects were hearing-impaired adults, also shown in Table 1.

| Table 1 about here |

II. **Accuracy.** Ideally, the differences between standard audiometry and thresholds measured using a new method would be similar to test-retest differences for standard audiometry alone. This would mean that a threshold measured using the new method would be just as accurate as measuring the standard audiometric threshold a second time. Average threshold difference data between iPad and audiometry therefore are compared to the test-retest data in the literature for standard audiometry; the iPad data are shown in the bottom row of Table 1. Eighty-nine percent of iPad vs. audiometer differences were within 10 dB. This compares with the audiometry test-retest
differences for hearing-impaired subjects (93%), but shows a larger spread than standard audiometry for normal subjects (99%). Ho et al. (2009) reported an average test-retest standard audiometry difference of 1.84 dB across all test frequencies for their hearing-impaired subjects, with SD = 6.68 dB. This compares closely with the average observed iPad/audiometry difference of 0.90 dB, with SD = 6.92 dB. The slightly larger variance of the iPad data when compared with normal subjects’ audiometric test-retest data is an expected outcome, given that the subjects in the iPad studies: 1) were much older than the subjects in the comparison studies; 2) received no verbal instruction on how to measure their thresholds; 3) were free to set their own threshold criteria in the iPad method; and 4) received no feedback of any kind during the test. Figure 3 shows the distribution of the iPad/audiometry threshold differences along with reported test-retest distribution for normal-hearing subjects of Schwanepoel et al., 2010. The similarity of the distributions is clear.

Home Version

The development objective for the Home Version of the hearing test was an accurate method of prescribing hearing device gain that does not require calibration. An accurate method is here defined as one that prescribes the same hearing device parameters as the NAL-NL2 parameters that would be prescribed from the user’s audiogram obtained via standard audiometry.

The major challenge in the Home Version is to get an estimate of the user’s audiometric configuration that is sufficiently accurate to result in accurate gain programming. This must be done in an environment in which it is impossible to calibrate -- that is, to know absolute sound pressure level in the user’s ear -- because the user can be using any commercially available sound card and earphones. The hi HealthInnovations approach (patent applied for) is to combine several minimally related sources of data about hearing loss to arrive at an accurate estimate of the user’s hearing. It is important to note that the Home Method does not measure an individual audiogram, and therefore it does not report an audiogram to the user. The purpose of the self-assessment is not to obtain an audiogram, but to program the hearing device so that it provides gain appropriate to the user’s hearing ability.

There are two main contributors to variance in audiometric configuration: severity and slope (Woods et al, 1999). Severity is estimated by taking advantage of the predictive relationship between the user’s score on a self-report questionnaire and his/her average pure tone thresholds. The slope is estimated by getting air conduction thresholds in each ear at 2 and 4 kHz using the modified method of adjustment described above. Although the absolute thresholds cannot be known in an uncalibrated system, the differences between thresholds at 2 and 4 kHz (slope) can be quantified, as can the differences across ears. (If the difference in measured thresholds across ears meets the referral criteria for asymmetry, then the user is referred for further evaluation and fitting.)

If the hearing is symmetrical, the method takes the information gathered from the user (questionnaire score, estimated average slope 2 – 4 kHz across ears, age, and gender), and
uses it to reconstruct the three most likely audiograms, based on demographic data from the
publicly-available National Health and Nutrition Examination Survey (Centers for Disease
Control and Prevention, National Center for Health Statistics). Each of the three memories of
the hearing device is then programmed with the NAL-NL2 parameters appropriate for each of
the three possible audiograms. The user is instructed to practice listening carefully with each
memory, and to choose the one that sounds the best to him/her.

**Evaluation of Gain Prescription Method.** The Hearing Screening Inventory (HSI) has been
reported to show good predictability for pure tone average (Coren & Hakstain, 1992). New
correlation data were collected using the questions from the HSI for use in Home method. At
each of two university locations (Northwestern University and University of Minnesota) 48 and
53 subjects, respectively, provided answers to the 12 questions of the HSI; they also had their
audiograms measured by an audiologist or gave permission for audiograms already on file at
the research location to be used. All research was approved by the IRB at the participating
institution.

The subjects who participated in the iPad studies at UWO and Arizona also answered the
questionnaire, so their data were combined with the data from the questionnaire-only subjects to
yield 144 persons with questionnaire results and audiograms (data from persons with
asymmetrical hearing losses were omitted). Figure 4 shows the measured relationship
between the HSI score and pure tone average (across ears) at four frequencies: 500, 1000,
2000, and 4000 Hz. The resulting linear regression equation is used in the Home Method to
predict the 4FPTA from the user’s questionnaire score; the known standard error is taken into
account in the estimation of loss severity.

Data from the same 144 subjects were used to evaluate the accuracy of the gain prescription
method to be used in the Home Version. We start with the assumption that for each of the 144
persons we know only four things:

1) questionnaire score (and therefore predicted 4FPTA, with known error range);

2) difference between 2 and 4 kHz thresholds each ear (with known error range attributable to
difference among earphones = ± 11 dB);

3) gender; and

4) age.

Using only the information that would be available from testing with the Home Version, we then
reconstruct the three most likely audiograms, and from them derive the NAL-NL2 parameters.
Of course, for these 144 subjects we also know the “true” audiograms, measured with standard
audiometric techniques, so we know the NAL-NL2 parameters that would have been prescribed
had their audiogram been known. We evaluate our method by comparing the NAL-NL2 gain
(for 65-dB SPL speech input) for the known audiogram with the gain that would be assigned by
the Home Method in the memory that is closest to the gain for the known audiogram.
Figure 5 shows the results. At all measurement frequencies the average difference between NAL-NL2 gain and Home Version gain is less than 1 dB, with standard deviations 2.6 – 3.9 dB. The overall rms error between audiogram and Home Version gain is 3.62 dB. In at least one memory, gain prescription is within 3 dB of NAL for 72% of users, and within 6 dB for 91% of users. This is well within the 3 – 9 dB that published data show as variance between initial settings for prescription methods and the gains that new users initially prefer (Marriage & Moore, 2004; Keidser et al., 2008).

Conclusions:

1. The Clinic Version of the method-of-adjustment self-test for measurement of AC thresholds (when using iPad in a calibrated and controlled environment) has been shown to:
   a. be as reliable as standard audiometry for hearing-impaired persons; and
   b. yield thresholds equivalent to thresholds measured via standard AC audiometry for hearing-impaired persons.
2. The hi HealthInnovations Home Version method for determining optimal gain prescription when calibration of earphones is not possible results in prescribed gains in at least one hearing device memory within 3-6 dB of NAL-NL2 gains for 91% of the 144 test subjects.
3. The hiHealthInnovations internet-based hearing test can be expected to:
   a. detect conditions requiring individualized services of a hearing health care professional, and make appropriate referrals for those services;
   b. prescribe appropriate gain for persons who are likely candidates for hi HealthInnovations open-fit hearing devices.

Future Improvements:

1. Device adjustments. Although: a) the Home and Clinic methods will result in accurate NAL-NL2 gain, and b) additional programs will bracket NAL gain in order to increase the odds that the user will find one of them satisfactory, it is still anticipated that many users will want to make further adjustments to device parameters to improve their listening satisfaction. An online expert system based on the one developed by Jenstad et al. (2003) will be implemented in the hi HealthInnovations website. Users will indicate which complaints are most descriptive of their experiences. Their hearing devices will then be re-programmed based on the expert system specifications.

2. Improvements to Home method. In the Clinic method, users provide answers to the questionnaire in addition to providing AC threshold data; in this environment the questionnaire answers are used to provide an accuracy check on the AC threshold data. The combined questionnaire and threshold data will also provide a rich source of data for continuous improvement of the accuracy of gain prescription in the Home method:
as more data are accumulated, computation methods in the Home method will be periodically refined.

Research Sponsorship and Disclaimer:

All research reported here was done at university sites under contract with Intricon Corporation. Data were collected independently at each site under controlled laboratory conditions for purposes of evaluating predictability of pure tone air conduction thresholds from self-tests and questionnaires. All data analysis, interpretation, and application to methods used in HealthInnovations programs were performed solely by the author. Universities that served as research sites do not necessarily endorse the conclusions stated herein, nor the HealthInnovations model for service delivery.
References


<table>
<thead>
<tr>
<th>Study</th>
<th>Method</th>
<th>Subjects</th>
<th>Percentage ± 10 dB</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Test-Retest</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Swanepoel et al. (2010)</td>
<td>Hughson-Westlake Standard Audiometry</td>
<td>30 normal-hearing</td>
<td>98.0%</td>
</tr>
<tr>
<td>Ho et al. (2009)</td>
<td>Hughson-Westlake Standard Audiometry</td>
<td>18 hearing-impaired</td>
<td>93.0%</td>
</tr>
<tr>
<td>iPad</td>
<td>Adjustment Self-Assessment</td>
<td>53 hearing-impaired</td>
<td>99.0%</td>
</tr>
<tr>
<td><strong>iPad vs. Standard Audiometry</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>iPad</td>
<td>Adjustment Self-Assessment</td>
<td>53 hearing-impaired</td>
<td>89.0%</td>
</tr>
</tbody>
</table>

Table 1: Test-Retest: Percent test-retest differences within ± 10 dB. iPad vs. Standard Audiometry: Percent threshold differences within ± 10 dB.
Figure 1. User interface for Flash and Air (non-touch-screen) versions of hi HealthInnovations hearing self-assessment. The volume slider is moved horizontally using a click-and-drag motion with a mouse.
Figure 2. User interface for iPod, iPad, and other tablet versions of hi HealthInnovations hearing self-assessment. The volume wheel is moved clockwise with the finger to increase volume, and counterclockwise to decrease volume.
Figure 3. Distribution of iPad vs. standard audiometry threshold differences. Comparison is estimated distribution of standard audiometry test-retest differences for normal-hearing subjects, taken from Table 3 of Swanepoel et al. (2010).
Figure 4. Binaural pure tone average (500, 1000, 2000, and 4000 Hz) air conduction thresholds as a function of score on Hearing Screening Inventory (Coren & Hakstain, 1994) for 144 subjects. All subjects had average thresholds 2000 – 6000 Hz within 15 dB across ears. Solid line is the best-fitting linear regression line; dashed lines are ±1 standard error.
Figure 5. Average assigned gain minus NAL-NL2 gain for 144 subjects with symmetric audiometric thresholds.