

Anne Elizabeth M. Javellana, MD
Ray U. Casile, MD

Department of Otolaryngology Head and Neck Surgery
St. Luke's Medical Center

The Correlation of Results between Pure Tone Audiometry and the Virtual Audiometer: A Simulated Air-Conduction Clinic-based Audiometer

ABSTRACT

The virtual audiometer (VA) is a software application that simulates a pure tone audiometer by delivering tones of different frequencies and intensities by air conduction.

Objectives: To determine correlation between hearing thresholds measured by virtual and pure tone audiometry and degree of agreement in their hearing loss classification.

Methods: Analytic, cross-sectional study set in a Philippine tertiary institution. Subjects were recruited from the outpatient department in September 2005 and comprised of cooperative, clinically normal and abnormal hearing individuals. Thirty-two subjects made 64 total ears tested (n=64), giving 0.90 correlation at 0.05 level of significance (α) ($p = 0.05$) and 99% power. Each subject underwent pure tone and virtual audiometry, the average thresholds calculated, and degree of hearing loss categorized according to classification by the World Health Organization (WHO). Results were not revealed until both examinations were completed. Data were stratified by frequency and compared by Pearson's correlation, while Kappa statistics determined degree of agreement between WHO grades of hearing impairment.

Results: Average age was 40.2 years, with 20-80 year range and 18.5 years standard deviation. Nine subjects were male while 23 were female (M:F ratio = 0.39). Of 64 ears, 34 were clinically normal and 30 were abnormal. Pearson's correlation demonstrated significant positive correlation between virtual and pure tone thresholds with 99% confidence at 0.05 level of significance. Kappa statistics also showed significant degrees of agreement in WHO grades by both instruments, meaning VA will probably categorize hearing loss in the same manner as pure tone audiometry.

Conclusion: A strong positive correlation exists between hearing thresholds measured by virtual and pure tone audiometry with a significant degree of agreement in hearing loss classification. This supports the possibility of using the virtual audiometer as a clinic-based, air-conduction audiometer for screening and monitoring. When used in conjunction with other examinations, valuable information on over-all integrity of the audiologic system may be ascertained.

Key words: *audiometer, audiometry, air-conduction, hearing test, hearing screening, hearing monitoring*

Correspondence:

Anne Elizabeth M. Javellana, MD
Department of Otolaryngology Head and Neck Surgery
St. Luke's Medical Center
279 E. Rodriguez Sr. Blvd.
Quezon City 1102
Philippines
Telefax: (632) 727 5543
Email: abjavellana@gmail.com
Reprints will not be available from the author.

No funding support was received for this study. The authors signed a disclosure that they have no proprietary or financial interest with any organization that may have a direct interest in the subject matter of this manuscript, or in any product cited in this report.

Presented at the Analytical Research Contest (2nd Place), Philippine Society of Otolaryngology Head and Neck Surgery 50th Annual Convention, EDSA Shangri-la Hotel, Pasig City, Philippines, December 2, 2006.

Pure tone audiometry is the current gold standard in evaluating auditory sensitivity. Regrettably, in developing countries like the Philippines, obtaining a pure tone audiogram may be difficult or problematic. There is want for a low-cost, quantitative, clinic-based tool for hearing screening and monitoring. Inexpensive tools, such as tuning forks, ballpen-click tests¹ and disability questionnaires² are available; however, these are either qualitative or subjective.

In an attempt to answer these needs, the virtual audiometer (VA) was developed by one of the authors, R.U. Casile, MD. It is a software application that simulates and delivers pure tones at different frequencies and intensities using a personal or laptop computer and ordinary earphones in a relatively quiet examination room. This allows the clinician to take quantitative measurements of auditory sensitivity in a clinic-based set-up.

Pure tone air-conduction thresholds measure the function of the total hearing system. In typical audiometric testing, pure tones in the octave range of 250-8000 Hz are presented to the listener by headphones or insert earphones. The Hughson-Westlake “ascending method,”³ is used to determine the hearing threshold – the lowest intensity at which a response is obtained three times. From this, the severity of the hearing loss is graded according to the World Health Organization (WHO) grades of hearing impairment⁴, presented in Table 1.

Table 1. World Health Organization Grades of Hearing Impairment (Geneva, 1991)¹⁰

Grade of Impairment	Corresponding audiometric ISO value (Average of 500, 1000 and 2000 Hz)
No impairment	25 dB or better (better ear)
Slight impairment	26 – 40 dB (better ear)
Moderate impairment	41 – 60 dB (better ear)
Severe impairment	61 – 80 dB (better ear)
Profound impairment including deafness	81 dB or greater (better ear)

VA approximates this determination by similarly presenting pure tones and quantifying threshold in terms of intensity. The difference is that the virtual audiometer is meant for clinic use, thereby, making it more accessible to both patients and physicians.

The requirements for VA examination are minimal. The examiner only needs a computer that runs on Windows™ operating system, intra-aural earphones, a relatively quiet room, and knowledge of his or her hearing threshold.

This study sought to determine (1) the correlation between the hearing threshold / hearing level (HL) measurements taken by virtual and pure tone audiometry and (2) the degree of agreement between the WHO grades of hearing impairment categorization of the average thresholds measured by the two examinations.

The null hypotheses were: (1) there is no correlation between hearing thresholds measured by virtual and pure tone audiometry; (2) there is no agreement between the WHO grades of hearing impairment categorization by the two examinations

The one-tailed alternate hypotheses were: (1) there is a positive correlation between hearing thresholds measured by virtual and pure tone audiometry; (2) there is a significant degree of agreement between the WHO grades of hearing impairment categorization by the two examinations.

MATERIALS AND METHODS

The principle of the virtual audiometer software is analogous to a tape recorder that records and collects the different tones of a real audiometer; that is, all frequencies from 500 Hz to 4000 Hz and the corresponding intensities of each tone from 0 dB to 100 dB. The pure tone audiometer from which the tones were recorded was a calibrated Beltone™ 10c (Beltone™, Chicago, IL, USA). The tones were captured using the built-in Sound Recorder in Microsoft Windows™ (Microsoft Corp., Redmond, WA, USA).

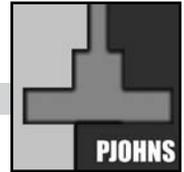
A software interface that would make the user/tester sound each tone was made using Corel™ Click and Create (now called Multimedia® Fusion Express, Clickteam, Paris, France). Simply stated, the interface retrieves the tone and intensity selected by the examiner as he or she clicks the frequencies and intensities marked on the screen, taking advantage of the speed of retrieval of the computer to sound off the desired tone.

This application claims priority from Philippine copyright registration number N 2007-19, registered on March 27, 2007.

To correlate results of virtual and pure tone audiometry, an analytic, cross-sectional study was set in a tertiary institution in Quezon City, Philippines in September 2005.

Assuming that hearing thresholds measured by virtual and pure tone audiometry have a correlation of at least 0.90, a level of significance (α) of 0.05 ($p = 0.05$) and a power of 99%, the estimated sample size was 24 subjects. The total number of individuals included was 32, with each ear tested separately, giving a total number of 64 tested ears ($n=64$).

From September 5 to 21, 2005, patients seen at the ENT



outpatient department of the authors' institution were enlisted in the study. Normal volunteers were also recruited from its staff and student population. History and physical examination ascertained whether subjects were clinically normal or abnormal, depending, respectively, on the absence or presence of otologic signs and symptoms (otorrhea, tympanic membrane perforation, impacted cerumen, subjective hearing loss, middle ear effusion, and so on). Inclusion of both clinically normal and abnormal subjects would allow measurement of a broader range of hearing levels and a more representative sample. After obtaining informed consent, all subjects underwent both virtual and pure tone audiometry.

Study subjects had to be (1) cooperative and (2) at least 7 years of age. Although conventional audiometry may be performed on children as early as 42-28 months of age³, the instructional set may need to be altered, and social reinforcement remains important. For this reason, small children may not be able to respond appropriately to the tone stimulus and, as such, were excluded. The following subjects were also excluded from the study: (1) uncooperative; (2) physically debilitated or unable to tolerate the procedure; (3) unable to follow or respond appropriately to instructions.

Primary data on hearing thresholds was collected directly from the subjects. Each underwent pure tone and virtual audiometric evaluation of both ears at the frequencies of 500, 1000, 2000 and 4000 Hz. The Hughson-Westlake method was employed for both examinations. Results were recorded in an audiogram, the average thresholds calculated, and the degree of hearing loss categorized according to the WHO classification. These allowed the collection of outcomes that were continuous (i.e., threshold measurements) and categorical (i.e., WHO classification). The results were not revealed until both examinations had been completed.

Pure tone audiometry was performed in a soundproof booth with a calibrated audiometer, while virtual audiometry was performed in a quiet room that was not acoustically dampened. This compared the virtual audiometer's performance to pure tone audiometry in the setting it was meant to be used. The set-up follows: testing was done in a relatively quiet environment. The more quiet the room, the better the set-up. Any possible noise source was eliminated (doors and windows closed, no outside conversations, towel placed over the computer housing). The authors tested the software in an air-conditioned room with ambient noise level measuring 60-63 dB SPL using a digital

display sound level meter (model number 33-2055, RadioShack Corp., TX, USA). A Pentium 2,23 3MHz laptop computer running on Windows 98™ (Microsoft Corp., Redmond, WA, USA) with the display screen facing the examiner was used. The software was loaded directly into the computer (but may be run from the CD copy). Intra-aural earphones, such as those from a portable music player, were connected to the speaker output jack of the computer. The authors used a pair of Sony™ intra-aural earphones (model MDR-E806, Sony Corp., Tokyo, Japan) with a frequency response of 20 Hz to 16 Khz and an impedance of 16 ohms. This type of headset comes packaged with most Sony Walkman™ portable music players (Sony Corp., Tokyo, Japan).

Calibration was performed at the onset of each exam. The examiner must have been previously tested by pure tone audiometry, knowing his or her threshold at 1000 Hz as the basis of calibration. With one earpiece on either the left or right ear, the examiner cupped hands over both ears to block out any extraneous noise. The signal intensity was matched to the examiner's threshold, that is, to a point where it was barely audible. This matched the hearing threshold measured by pure tone audiometry, calibrating the software.

The earphones were then transferred to the examinee, one earpiece at a time, testing each ear separately. The examinee also cupped his or her hands over both ears to further isolate the signal (alternatively, ear protectors/muffs may be positioned over the earphones). A test frequency was selected and the examiner determined the threshold using the Hughson-Westlake method. The signal was presented well above the threshold, decreased to a level of inaudibility in 10- to 15-dB steps, and then increased in "up 5-dB, down 10-dB steps" until the single hearing level at which a response was obtained three times was reached. The same was done for the remaining test frequencies. The earpiece was transferred to the other ear and the procedure repeated.

At the end of each examination, the thresholds of both ears at all four frequencies were plotted in an audiogram. The average thresholds were calculated and the WHO grade assigned. The procedure was repeated on each examinee.

Results were correlated at each test frequency. A single subject's VA and pure tone audiometry results were compared to each other, minimizing inter-individual variation. Virtual and pure tone measurements were presented side-by-side and stratified according to (1) frequency and (2) WHO classification. The degree of correlation between VA and pure tone threshold measurements at each frequency was analyzed using the Pearson

Product Moment Correlation, while the degree of agreement between the WHO classes were analyzed using the Kappa statistic. Both tests were conducted at 0.05 level of significance. All data were encoded and analyzed using SPSS (Statistical Package for the Social Sciences) v.14 statistical software (SPSS, Inc., Chicago, IL, USA).

RESULTS

Thirty-two subjects were enrolled in the study, giving a total of 64 tested ears (N=64). The average age was 40.2 years, with a 20-80 year range and a standard deviation of 18.5 years. Nine subjects were male and 23 were female (M:F ratio = 0.39). Of the 64 ears tested, 34 were clinically normal and 30 were clinically abnormal or presented with ear symptoms.

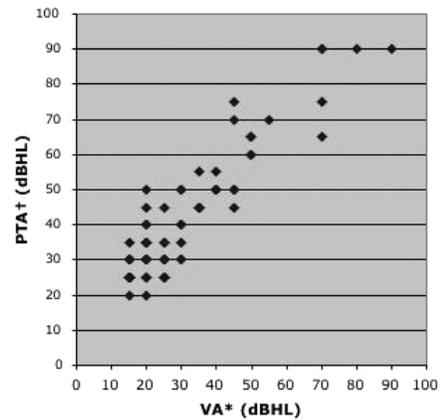
A significant positive correlation was found between virtual and pure tone audiometric thresholds on all tested frequencies and the average threshold (Figures 1 to 5). Pearson’s correlation factors were greater than zero and approximate one, with *p* values (Sig.) <0.01, indicating a direct relationship between the results with at least 99% confidence.

In addition, a significant degree of agreement was found between the WHO grading arrived at by the two examinations. Of the 39 subjects found to have normal hearing by VA, 28 (71.8%) also had normal hearing by pure tone audiometry, while 11 (28.2%) had mild hearing loss. Of the 7 subjects found to have mild hearing loss by VA, pure tone audiometry confirmed 4 (57.1%), while 3 (42.9%) actually had moderate hearing loss. Of the 12 subjects determined to have moderate hearing loss by VA, 9 (75.0%) also had moderate hearing loss by pure tone audiometry, while 3 (25.0%) had severe hearing loss. Of the 5 subjects graded as severe by VA, pure tone audiometry confirmed 4 (80.0%) and determined 1 (20.0%) to be profound. The single subject graded as profound by VA had severe hearing loss by pure tone audiometry.

The *p* values of the kappa statistics for all five WHO classifications were less than 0.01, meaning a subject determined to have a particular grade of hearing loss by pure tone audiometry would probably be categorized by VA in the same manner.

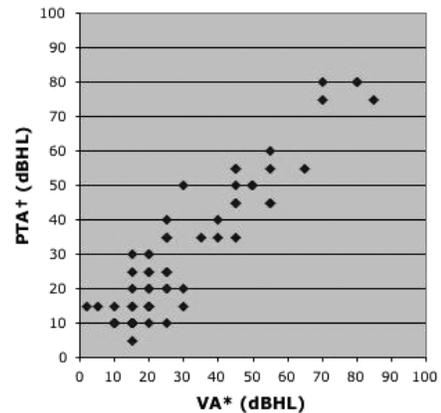
The virtual audiometer’s specificity and positive predictive value were both 100%. Sensitivity was 69.44%, while negative predictive value was 71.8%.

Figure 1. Hearing Thresholds Measured at 500Hz



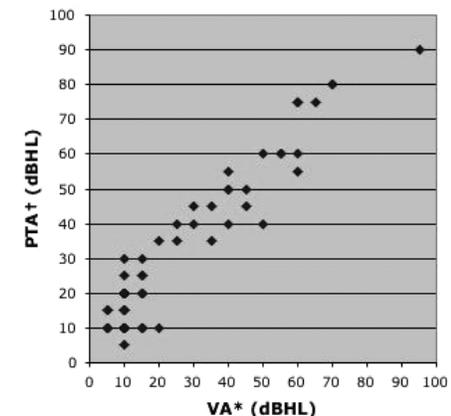
* VA - hearing thresholds in dBHL measured by the virtual audiometer
† PTA - hearing thresholds in dBHL as measured by the pure tone audiometer

Figure 2. Hearing Thresholds Measured at 1000 Hz



* VA - hearing thresholds in dBHL measured by the virtual audiometer
† PTA - hearing thresholds in dBHL as measured by the pure tone audiometer

Figure 3. Hearing Thresholds Measured at 2000 Hz



* VA - hearing thresholds in dBHL measured by the virtual audiometer
† PTA - hearing thresholds in dBHL as measured by the pure tone audiometer

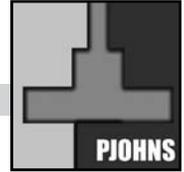
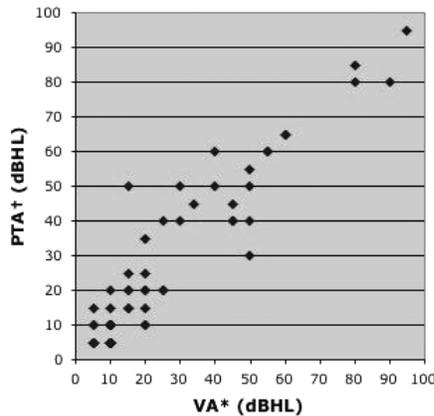
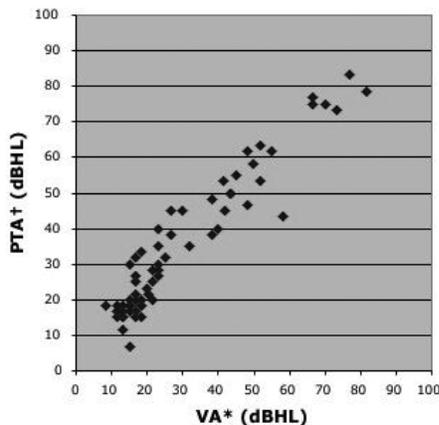


Figure 4. Hearing Thresholds Measured at 4000 Hz



* VA - hearing thresholds in dBHL measured by the virtual audiometer
† PTA - hearing thresholds in dBHL as measured by the pure tone audiometer

Figure 5. Average Threshold Measurements



* VA - hearing thresholds in dBHL measured by the virtual audiometer
† PTA - hearing thresholds in dBHL as measured by the pure tone audiometer

DISCUSSION

Pure-tone audiometry is currently the gold standard in evaluating auditory sensitivity.³ Pure-tone signals are produced by an oscillator and delivered through air and bone conduction. The American National Standards Institute (ANSI) S3, 20-1973 defines the threshold of audibility as "the minimum effective sound pressure level of an acoustic signal producing an auditory sensation 'in a specified fraction of the trials.'" Often, threshold is defined as the lowest signal intensity at which multiple presentations are detected 50% of the time.³ Clinically,

audiometric thresholds are displayed on an audiogram. Data is presented in hearing levels (HL) which are calibrated to referent sound pressures (ANSI S3, 6-1969, 1970) that represent the hearing sensitivity of normal, young adults when tested under reasonably quiet test conditions. The pure tone audiometer requires a soundproof booth, regulation headphones, a bone transducer and regular calibration to ensure accurate measurements.

The virtual audiometer has several contrasting features. First, it does not have an oscillator. The different tone frequencies were recorded from a pure tone audiometer and stored into a database, then reproduced at different intensities by an authoring software. It is in this sense that this audiometer is "virtual." It does not produce the pure tones itself, but simulates them. Difficulties in reproducing the pure tones were encountered in sound pressure levels above 80 decibels. At this point, the waveforms were clipped and no longer the sine waves of pure tones, but square waves with distortion. Below 80dB, the tones were accurately simulated.

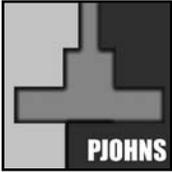
Second, since its aim is screening and monitoring, only air conduction is measured. Accordingly, bone transducers are not employed. The goal is to keep its requirements simple: a personal computer equipped with Windows™ operating system and a sound card, intra-aural earphones and knowledge of one's own threshold level. The examiner himself calibrates the virtual audiometer. These also seek to minimize the cost of the examination.

Third, in order to increase accessibility, VA was designed to be used in the clinics, not in a sound-insulated room.

In spite of these differences, favorable results were still achieved. Pearson's correlation showed a strong positive correlation at all four test frequencies and for the average hearing thresholds. In addition, the Kappa statistics also showed a significant degree of agreement across all WHO classes, strengthening the inference of Pearson's correlation and, consequently, the results of the study.

In conclusion, this study established a strong positive correlation between hearing threshold measured by virtual and pure tone audiometry. Furthermore, there was a significant degree of agreement in WHO grading of hearing loss between the two examinations.

The results of this study support the possibility of using the virtual audiometer as a clinic-based, air-conduction audiometer for screening and monitoring patients. When used in conjunction with other clinic examinations, valuable information on the



overall integrity of the subject's audiologic system may be ascertained.

In order to decrease instrumentation errors, the software itself should be fine-tuned to ensure accurate reproduction of pure tones and its calibration may be extended to all test frequencies. In addition, different types of earphones can be tested with the software. An MP3 format of the program is being developed by the authors, making the software more portable by using handy MP3 players. Future studies with increased sample size may attain a more representative cross-sectional population, stratifying sampling into all five WHO classes. Statistical analysis may be further stratified according to WHO classification, thereby allowing more subtle relationships to manifest.

ACKNOWLEDGEMENTS

We gratefully acknowledge Isaac David E. Ampil, II, MD, MSc, Co-director for Clinical Research, Research and Biotechnology Division, St. Luke's Medical Center for statistical analysis and content organization.

REFERENCES

1. Martinez NV. Prevention of Deafness Program in the Philippines. IFHOH [Internet]. Available from: <http://www.ifhoh.org/papers/filipino-2.htm>
2. Ventry IM, Weinstein BE. The hearing handicap inventory for the elderly: a new tool. *Ear Hear*. 1982 3(3):128-134. May-Jun.
3. Cummings CW, Flint PW, Harker LA, Haughey BH, Richardson MA, Robbins KT, Schuller DE, Thomas JR, editors. Cummings otolaryngology-head and neck surgery. 4th ed. Vol. 4. Philadelphia (PA): Mosby, c2005. 3374 p.
4. Informal working group on prevention of deafness and hearing impairment. Report of the informal working group on prevention of deafness and hearing impairment programme planning. Geneva (Switzerland): World Health Organization; 1991 Jun.
5. Newman CW, Sandrigde SA. Hearing loss is often undiscovered, but screening is easy. *Cleve Clin J Med* [Internet]. 2004 Mar; 71(3): 225-232. Available from: www.ccm.org/PDFFILES/Newman304.pdf.
6. Yueh B, Shapiro N, MacLean CH, Shekelle PG. Screening and management of adult hearing loss in primary care. *JAMA*. 2003; 289:1976-1985.
7. Cienkowski KM. An evaluation of the comparative hearing test. *Internet J Geriatr Gerontol* [Internet]. 2003; 1(1):[about 9 screens]. Available from: <http://www.ispub.com/ostia/index.php?xmlFilePath=journals/ijgg/vol1n1/hearing.xml>
8. Stuart A, Stenstrom R, Tompkins C, Vandenhoff S. Test-retest variability in audiometric threshold with supraaural and insert earphones among children and adults. *Audiology*. 1991;30(2):82-90.
9. Landry JA, Green WB. Pure-tone audiometric threshold test-retest variability in young and elderly adults. *J Speech Lang Pathol Audiol* [Internet]. 1999 Jun; 23(2):74-80. Available from: http://www.caslapa.ca/PDF/seniors%20articles/puretone_audiometric_thresholdtest_elderly.pdf