International Journal of Audiology 2012; 51: 606–610

informa
healthcare

Original Article

Clinical evaluation of a computerized self-administered hearing test

Anna Chi Shan Kam*,†, John Ka Keung Sung*,†, Tan Lee‡, Terence Ka Cheong Wong*,† & Andrew van Hasselt*,†

*Department of Otorhinolaryngology, Head and Neck Surgery, Chinese University of Hong Kong, Shatin, Hong Kong, †Institute of Human Communicative Research, Chinese University of Hong Kong, Shatin, Hong Kong, and ‡Department of Electronic Engineering, Chinese University of Hong Kong, Shatin, Hong Kong

Abstract

Objective: To establish the reliability and validity of a computerized self-administered hearing test. *Design:* Cross-sectional within a comparative study of subjects. *Study sample:* Subjects were 100 Chinese adults who attended the audiology clinic in a hospital for a hearing test. *Results:* There was no significant difference in the thresholds of unmasked air-conduction hearing obtained with the computerized self-administered hearing test via a smartphone and those obtained with standard pure-tone audiometry. High test-retest reliability was observed with the self-administered hearing test (intraclass correlation coefficient = 0.95), and was comparable with that observed in standard pure-tone audiometry (intraclass correlation coefficient = 0.97). The thresholds of the self-administered hearing test measured in a sound-proof booth were not significantly different from those measured in a quiet office room. *Conclusions:* The results suggest that the computerized self-administered hearing test is a reliable and valid measure of unmasked air-conduction hearing thresholds.

Key Words: Pure-tone audiometry; self-administered; hearing loss

Hearing loss is an important public health problem because of its disabling consequences. Hearing is essential for daily communication and social interaction. Depending on its degree of severity, hearing impairment may negatively affect physical, cognitive, behavioral, and social functions, work productivity, and even general quality of life. Hearing loss has been estimated to be the most prevalent disability in developed countries (Davis, 1997). In 2003 - 04, 16.1% of US adults (29 million Americans, aged 20–69 years) had speech-frequency (500, 1000, 2000, and 4000 Hz) hearing loss (Agrawal et al, 2008). About 16% of UK (Davis, 1989) and 17% of Italian adults (17–80 years old) (Quaranta et al, 1996) have a ≥25-dB hearing level (HL) impairment in both ears or in the better ear. Many of these people are unaware that they have suffered a decrease in hearing sensitivity. Unfortunately, few people undergo regular or frequent hearing tests as a part of their routine health care, partly because of the lack of a simple, convenient, and inexpensive hearing test (Margolis & Morgan, 2008), and some may not have even thought about taking a hearing test in their life-time. Failure to recognize the hearing loss and institute appropriate intervention may worsen its impact (Brink & Stones, 2007; Chia et al, 2007; Dalton et al, 2003). The economic, social and educational opportunity loss, due to unattended hearing loss, can be reduced or avoided by early detection and intervention.

Pure-tone audiometry is by far the most common form of hearing measurement. It is typically conducted by audiologists or trained health care professionals in a sound booth. However, such audiological services may not be readily accessible in areas where professional staff and/or audiological equipment are not available. Most patients who undergo pure-tone audiometry should be able to follow instructions and provide responses to automated procedures (Margolis & Morgan, 2008). By implementing an automated self-administered hearing test, more patient-contact time could be made available for professional staff to conduct sophisticated tests and manage patients who are difficult to test (Margolis & Morgan, 2008). Recently, many studies have been carried out to validate automated audiometry (Ho et al, 2009; Ishak et al, 2011; Margolis et al, 2010, 2011; Margolis & Moore, 2011). However, in most cases, such automated tests did not alleviate the problem of availability of audiological equipment and test settings (typically a sound booth or sound-treated room).

Another way of increasing accessibility of hearing tests to the general public is to use personal computers and smartphones as the test media. With the vigorous growth in penetration rate of smartphones, automated hearing tests in the form of mobile apps have become more popular. However there is a lack of validation and data supporting their use (Coleman, 2011). Our research team has developed a computerized self-administered hearing test which can be used with everyday electronic audio devices such as computers and smartphones. The aims of this investigation were to provide an easier and more accessible means for people to take a valid hearing test, so that they can learn more about their hearing status and be alerted when it is time to seek professional advice for their hearing problem. This study, therefore, investigated the validity and reliability of the

Correspondence: Anna Chi Shan Kam, Department of Otorhinolaryngology, Head and Neck Surgery, The Chinese University of Hong Kong, 6/F, Clinical Sciences Building, Prince of Wales Hospital, Shatin, N.T., Hong Kong. E-mail: annakam@ent.cuhk.edu.hk

(Received 16 January 2012; accepted 18 April 2012)

DOI: 10.3109/14992027.2012.688144

computerized self-administered hearing test in a sound booth and the feasibility of conducting such tests in non-sound-proof settings.

Materials and Methods

Procedures

All potential subjects first completed a routine audiological assessment in a sound-proof booth at the Audiology Clinic of the Prince of Wales Hospital, where standard pure-tone audiometry and an otoscopic examination were performed. The audiometry thresholds were measured according to the procedures recommended by the British Society of Audiology (British Society of Audiology, 2004). Air-conduction thresholds were obtained at 250, 500, 1000, 2000, 4000, and 8000 Hz, and bone-conduction thresholds at 500, 1000, 2000, and 4000 Hz were also obtained using a Grason-Stadler GSI 61 clinical audiometer with Telephonics TDH-50 headphones. All transducers were calibrated in accordance with ANSI S3.6-2004 (American National Standards Institute, 2004). Subjects were invited to participate in the study if the following inclusion criteria were met. These were: (1) age over 18 years; (2) air-conduction thresholds at all octave frequencies tested (250–8000 Hz) were within the range of the self-administered hearing test, that is, with no more than 55 dB HL loss; and (3) good oral communication abilities and the ability to understand and give informed consent. The subjects then attended a 30-minute test session in a sound booth. After obtaining the informed consent, a history form including demographic information was completed. The subjects then took the computerized self-administered hearing test. A subgroup of 20% of the recruited subjects (n = 20) was selected at random to assess test-retest reliability, and another subgroup (n = 20) was selected at random to repeat the self-administered hearing test in a quiet room, rather than the sound booth. The quiet room was an office room in the audiology clinic with an ambient noise level of 41 dBA. Before each retest, both for the test-retest reliability and test environment effect investigation, the subjects were required to remove the earphones and put them in again. An investigator was always available in the test room for enquiry and monitoring.

Computerized self-administered hearing test

A computerized self-administered hearing test for mobile devices has been developed. In this study, the hearing test was administered on a smartphone (Apple iPhone 3GS with iOS4), and standard earphones which come with the basic model of an Apple iPod were used as the transducer for testing. The earphones used are shown in Figure 1. There is no standard for calibrating the output of earphones used with audio devices. The earphone was plugged into a KEMAR pinna which attached to a Zwislocki coupler. The other side of the coupler was connected to a Bruel & Kjær Sound Level Analyser (Type 2260 Investigator) which measured the output of the system via the earphone. The right and left earphone was measured separately. At each frequency tested, the signal amplitude was adjusted according to the readings from the sound level meter so as to ensure that all output levels (in 5-dB steps) measured in dB SPL were equivalent to the dB HL values set in the system with the conversion from the Zwislocki coupler to 2-cc (HA-1) SPL (Wilber et al, 1988) were applied. Six test frequencies were used (250, 500, 1000, 2000, 4000, and 8000 Hz). Due to the output limit (without distortion) of the earphones, the test ranges at different frequencies varied from −10 to 55 dB HL to −10 to 70 dB HL, as shown in Table 1. Pulsed pure tones of all test frequencies were presented on the right side

Figure 1. Photo of the earphones used in the study.

first. The sequence of test frequency followed that of conventional pure-tone audiometry. The subject was required to press the '−' sign if the test tones could be heard, and, when the test tones disappeared, the subject had to press the '+' sign once. By pressing the '+' or '−' sign, the test signal would be increased or decreased by 5 dB. The subject could repeat the steps as many times as they needed until the test signal could just be heard. This level of volume would be set as the threshold and be saved by the program once the subject pressed the 'OK' button on the touch screen of the smartphone. Figure 2 shows a screenshot of the hearing test on a smartphone.

Method of adjustment

Verbal instructions were given before the test and the subjects were reminded to firmly position the right and left earphones on the respective ears with a diagram as shown in Figure 3 being displayed on the device. The following verbal instructions were given, "Please put on the earphones as tight as possible on the correct side. Please adjust the volume level using the '+' or '−' button so you can just barely hear the tone. Press the '+' button if you cannot hear the tone and press the '−' button if you can hear the tone. You can repeat the steps as many times as you like until the tone could just be heard. Please press 'OK' when you are done."

Ethical considerations

Local research ethics approval was granted for the study.

Results

Subject demographics

One hundred adult patients, who attended the Audiology Centre of the Prince of Wales Hospital for pure-tone audiometry and met all

Table 1. Test range of the computerized self-administered hearing test via a smartphone.

	Frequency (Hz)					
Test range (dB HL)	*250*	*500*	*1000*	*2000*	*4000*	*8000*
Minimum	−10	−10	−10	−10	−10	−10
Maximum	55	60	65	70	70	60

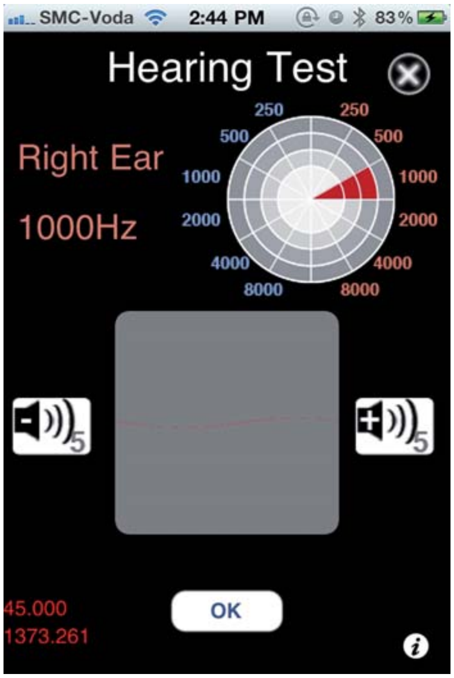

Figure 2. Screenshot of the hearing test application on a smartphone.

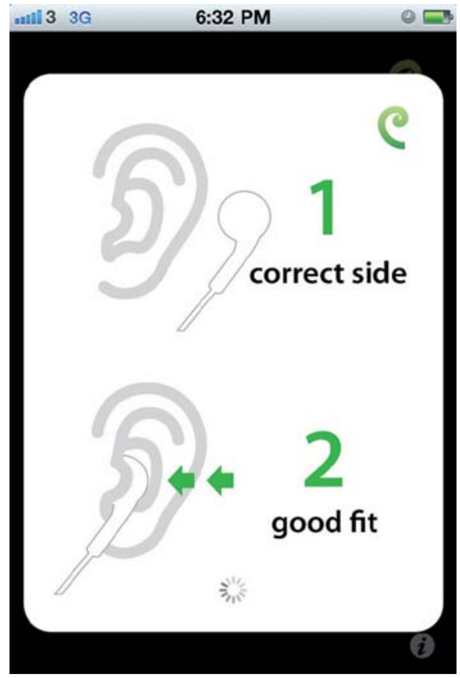

Figure 3. Instruction screen display for subjects before testing.

the inclusion criteria, participated in the study. The subjects were aged from 18 to 73 (mean = 47.31, standard deviation [SD] = 13.39) and included 59 women and 41 men. Fifty-eight (58%) of the subjects had normal hearing bilaterally with pure-tone thresholds equal to or better than 25 dB HL (American National Standards Institute, 1989) across all octave frequencies at 250–8000 Hz. Table 2 shows the demographic data of the subjects.

Measured unmasked air-conduction thresholds

The average hearing threshold across both ears at all of the six frequencies tested was 21.5 dB HL (SD = 11.8). The means and SDs of thresholds at each frequency and ear obtained using the two methods are listed in Table 3. When 1200 paired (100 subjects × 2 ears × 6 frequencies) air-conduction thresholds obtained by pure-tone audiometry and the self-administered hearing test were compared, 26% of the paired thresholds had 100% agreement. Eighty-eight per cent of the paired thresholds fell within 10 dB of each other. Figure 4 shows the distribution of the differences between thresholds obtained from pure-tone audiometry and the self-administered hearing test at each ear and frequency.

Mean differences (self-administered hearing test threshold minus pure-tone audiometric threshold) and mean absolute differences in the thresholds for each frequency and averaged across all frequencies are shown in Table 4. Averaged across the right and left ears, the differences ranged from −4.3 dB at 1000 Hz for the right ear to 7.6 dB at 8000 Hz for the left ear. The mean difference for all ears and frequencies was 0.5 dB, suggesting that there were no significant differences between self-administered hearing test thresholds and pure-tone audiometric thresholds. Absolute differences averaged across the right and left ears ranged from 4.7 dB at 2000 Hz to 9.6 dB at 8000 Hz. Across all ears and frequencies, the mean absolute difference was 6.3 dB.

A three-way analysis of variance (ANOVA) with repeated measures on three factors (ear, test frequency, and test method) was performed. The 'test method' factor was insignificant, indicating that the thresholds obtained were not significantly different between the two test methods. The 'ear' factor was insignificant, indicating that the thresholds obtained were not significantly different between the right and left ear. However, the 'frequency' factor was significant ($F = 36.87$; $p < 0.01$) which could be explained by the generally higher thresholds at high frequencies than at low frequencies, especially in subjects with hearing loss.

After controlling for age and hearing loss, a further ANOVA was performed on the level of education of the subjects on threshold determination. No significant effect was observed, indicating that the validity of the computerized self-administered hearing test was not affected by the subjects' level of education.

Test-retest reliability

Demographic data on the 20 subjects who participated in this part of study are listed in Table 2. The intra-session test-retest correlations (intraclass correlation coefficients) were 0.97 and 0.95 ($p < 0.01$) for the pure-tone audiometric thresholds and the self-administered hearing test thresholds, respectively, which indicated good test-retest reliability. Table 5 shows the percentages of participants who had test-retest hearing threshold variability of no more than 5 dB and no more than 10 dB, respectively, at each frequency and each measurement. Variability for pure-tone audiometry was within 5 dB for 95–100% of ears in the frequency range from 500 to 4000 Hz, and was within 10 dB for 100% of ears in all frequencies tested. Variability for the self-administered hearing test was within 5 dB for 89–100% of ears in the frequency range of 500 to 8000 Hz, decreasing to 83% at 250 Hz, and was within 10 dB for 92–100% of ears in all frequencies tested.

Sound booth versus quiet room

Demographic data on the 20 subjects who participated in this part of study are listed in Table 2. A three-way ANOVA with repeated

Table 2. Demographic data of the study subjects.

	All subjects	*Subgroup 1*	*Subgroup 2*
Age (years)			
Range	18–73	21–62	20–56
Mean	47	48	46
SD	13	9	11
Gender			
Male	41	9	7
Female	59	11	13
Education level			
None	2	0	0
Primary	19	4	3
Secondary	51	11	13
Tertiary or above	28	5	4
Type of loss (number of ears)			
Normal	116	19	22
Sensorineural	60	18	11
Conductive	13	2	5
Mixed	11	1	2
Total number of ears tested	200	40	40

SD, standard deviation.

measures on three factors (test environment, test frequency, and test ear) was performed. The 'test environment' factor was insignificant, indicating that the thresholds obtained with the self-administered hearing test in a quiet room were not significantly different from those obtained in a sound booth. The 'test frequency' factor was significant ($F = 12.11$; $p < 0.01$), which could be explained by the generally increased thresholds at high frequencies than at low frequencies, especially in subjects with hearing loss. The 'test ear' factor was insignificant, indicating that the thresholds obtained were not significantly different between the right and left ear.

Discussion

There was no significant difference in the unmasked air-conduction hearing thresholds obtained from the computerized self-administered hearing test using a smartphone and those obtained with standard pure-tone audiometry that was administered by audiologists with clinical audiometers. This showed that the computerized self-administered hearing test is as accurate as standard pure-tone audiometry. High test-retest reliability was observed with the computerized self-administered hearing test (intraclass correlation coefficients = 0.95), which was comparable with that of standard pure-tone audiometry (intraclass correlation coefficient = 0.97). There was no significant difference in the thresholds obtained with the computerized self-administered hearing test in a quiet room and those obtained in a sound-proof booth. This means that, if the ambient noise level is low enough (41 dBA in this study), the results from the self-administered hearing test in a quiet room could be as accurate as those obtained in a sound booth.

Table 3. Mean thresholds obtained at each frequency and ear via two test methods, with associated standard deviations (SD).

		Frequency (Hz)						
		250	*500*	*1000*	*2000*	*4000*	*8000*	*All*
Pure-tone audiometry								
Right	Mean	19.8	19.7	21.5	19.3	22.6	26.1	21.5
	SD	10.8	9.9	10.7	10.9	13.6	15.0	11.8
Left	Mean	20.2	19.1	19.6	18.6	23.5	26.2	21.2
	SD	11.3	10.5	9.6	12.6	12.7	15.9	12.1
Both	Mean	19.8	19.7	21.5	19.3	22.6	26.1	21.5
	SD	10.8	9.9	10.7	10.9	13.6	15.0	11.8
Self-administered hearing test								
Right	Mean	18.0	19.1	17.2	21.9	21.1	26.7	20.7
	SD	11.3	9.7	9.6	9.3	11.0	14.8	11.0
Left	Mean	21.3	21.0	17.6	22.5	21.7	33.8	23.0
	SD	11.7	9.8	8.1	11.4	10.9	14.6	11.1
Both	Mean	18.0	19.1	17.2	21.9	21.1	26.7	20.7
	SD	11.3	9.7	9.6	9.3	11.0	14.8	11.0

Comparison with other automated hearing tests and mobile apps

Previous studies have shown that automated diagnostic audiometry, including bone-conduction and masking, is valid and reliable (Ho et al, 2009; Ishak et al, 2011; Margolis et al, 2010, 2011; Margolis & Moore 2011). Average absolute differences in air-conduction threshold across all ears and frequencies between automated test and manual pure-tone audiometry of 2.4 dB (Swanepoel et al, 2010) and 3.6 dB (Margolis et al, 2010) have been reported. The average absolute difference obtained in the present study was 6.3 dB. The higher variability with the computerized self-administered hearing test might be explained by the use of a non-clinical transducer.

Recently many hearing-related apps have become available for free or at a very low cost (Coleman 2011). Some of them just provide a hearing test while some include speech-in-noise assessment or even amplification function. However none of these have been validated and no available data could be found for comparison with the present study.

Limitation of the self-administered hearing test

Due to the output limit (without distortion) of the earphones of the smartphones used in the study, levels of hearing impairment worse than a moderate grade could not be tested using the current version of the application. The accuracy of the test was investigated in quiet only in the present study, thus a quiet test environment is recommended. Since the self-administered hearing test could only be used to obtain unmasked air-conduction hearing levels, hearing loss of

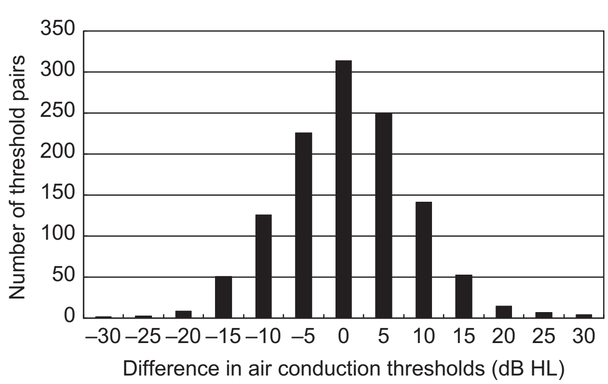

Figure 4. Distribution of the differences between thresholds obtained from pure-tone audiometry and the self-administered hearing test for each ear and frequency. Note. The difference is calculated by subtracting the threshold obtained with pure-tone audiometry from those obtained with the self-administered hearing test.

Table 4. Mean differences and absolute differences between self-administered hearing test thresholds and pure-tone audiometric thresholds, with associated standard deviations (SD).

		Frequency (Hz)						
		250	*500*	*1000*	*2000*	*4000*	*8000*	*All*
Average difference								
Right	Mean	−1.8	−0.7	−4.3	2.9	−1.6	0.7	−0.8
	SD	9.4	7.8	6.6	6.1	7.0	7.4	7.4
Left	Mean	1.2	1.9	−2.0	3.9	−1.8	7.6	1.8
	SD	9.6	7.8	5.9	7.1	7.5	9.5	7.9
Both	Mean	−0.3	0.6	−3.1	3.3	−1.7	4.1	0.5
	SD	8.6	6.9	5.3	5.9	6.4	7.0	6.7
Average absolute difference								
Right	Mean	7.2	5.9	5.9	4.7	5.7	5.3	5.8
	SD	6.3	5.2	5.2	4.7	4.4	5.2	5.2
Left	Mean	8.1	6.4	4.7	6.1	5.9	9.6	6.8
	SD	5.4	4.9	4.2	5.3	5.0	7.5	5.4
Both	Mean	7.6	6.1	5.3	5.4	5.8	7.4	6.3
	SD	4.7	4.1	3.5	4.4	3.7	4.7	4.2

the worse ear in people with asymmetrical hearing may be underestimated. Unlike standard pure-tone audiometry which uses standardized insert earphones or headphones with calibrated headband force, the self-administered hearing test uses general smartphone earphones which may be plugged loosely into the ear. This may result in a large variability in the thresholds obtained.

Application of the self-administered hearing test

The computerized self-administered hearing test was developed to provide a simple, convenient means for people to take a hearing test. This study showed that it can be used by people of different ages and education level, which means that the test could be readily applied to the general population. It is not designed to replace standard pure-tone audiometry which is performed by an audiologist in a sound booth. However, due to its accessibility via smartphones and other computer devices, people can easily have their hearing tested, and be alerted to seek professional advice if their test results are abnormal. Thus, the computerized self-administered hearing test can be a convenient means to raise public awareness on hearing health.

Conclusion

This study showed that the procedures of the computerized self-administered hearing test and the use of a smartphone to obtain unmasked air-conduction thresholds are valid and reliable.

Table 5. Percentages of ears that had variability ≤5 dB and ≤10 dB for each test and frequency.

	Frequency (Hz)					
Test-retest variability	*250*	*500*	*1000*	*2000*	*4000*	*8000*
Pure-tone audiometry						
≤5 dB (%)	85	100	100	100	95	90
≤10 dB (%)	100	100	100	100	100	100
Self-administered hearing test						
≤5 dB (%)	83	89	97	92	100	94
≤10 dB (%)	94	92	100	97	100	100

Acknowledgements

The authors gratefully thank Mr. Paul Lee and Mr. Andy Cheng from Ximplar Limited for the technical support in software development, and Miss Zoe Chan for her participation in the data collection.

Declaration of interest: This study was supported by the Direct Grant for Research from the Faculty of Medicine, the Chinese University of Hong Kong. The development of the self-administered hearing test application was supported by the Downstream Development Seed Fund from the Chinese University of Hong Kong. The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References

Agrawal Y., Platz E.A. & Niparko J.K. 2008. Prevalence of hearing loss and differences by demographic characteristics among US adults: Data from the National Health and Nutrition Examination Survey, 1999–2004. *Arch Intern Med*, 168, 1522–1530.

American National Standards Institute. 1989. *Specification for audiometers (ANSI S3.6 - 1989)*. New York: ANSI.

American National Standards Institute. 2004. *Specification for Audiometers (ANSI S3.6 - 2004).* New York: ANSI.

Brink P. & Stones M. 2007. Examination of the relationship among hearing impairment, linguistic communication, mood, and social engagement of residents in complex continuing-care facilities. *Gerontologist*, 47, 633–641.

British Society of Audiology. 2004. Recommended procedure: Pure tone air and bone conduction threshold audiometry with and without masking and determination of uncomfortable loudness levels.

Chia E., Wang J.J., Rochtchina E., Cumming R.R., Newall P. et al. 2007. Hearing impairment and health-related quality of life: The Blue Mountains hearing study. *Ear Hear*, 28, 187–195.

Coleman M. 2011, There's a hearing app for that. *Hear J*, 64(11), 12–16.

Dalton D.S., Cruickshanks K.J., Klein B.E.K., Klein R., Wiley T.L. et al. 2003. The impact of hearing loss on quality of life in older adults. *Gerontologist*, 43, 661–668.

Davis A.C. 1997. Epidemology of hearing disorders. In: A.G. Kerr (ed.) Scott Brown's Otolaryngology. Boston: Butterworth-Heineman.

Davis A.C. 1989. The prevalence of hearing impairment and reported hearing disability among adults in Great Britain. *Int J Epidemiol*, 18, 911–917.

Ho A.T.P., Hildreth A.J. & Lindsey L. 2009. Computer-assisted audiometry versus manual audiometry. *Otol Neurotol*, 30, 876–883.

Ishak W.S., Zhao F., Stephens D., Culling J., Bai Z. et al. 2011. Test-retest reliability and validity of Audioscan and Békésy compared with pure tone audiometry. *Audiol Med*, 9, 40–46.

Margolis R.H., Frisina R. & Walton J.P. 2011. AMTAS®: Automated method for testing auditory sensitivity: II. Air conduction audiograms in children and adults. *Int J Audiol*, 50, 434–439.

Margolis R.H., Glasberg B.R., Creeke S. & Moore B.C.J. 2010. AMTAS®: Automated method for testing auditory sensitivity: Validation studies. *Int J Audiol*, 49, 185–194.

Margolis R.H. & Moore B.C.J. 2011. AMTAS®: Automated method for testing auditory sensitivity: III. Sensorineural hearing loss and air-bone gaps. *Int J Audiol*, 50, 440–447.

Margolis R.H. & Morgan D.E. 2008. Automated pure-tone audiometry: An analysis of capacity, need, and benefit. *Am J Audiol*, 17, 109–113.

Quaranta A., Asennato G. & Sallustio V. 1996. Epidemiology of hearing problems among adults in Italy. *Scand Audiol Suppl*, 25, 9–13.

Swanepoel W., Mngemane S., Molemong S., Mkwanazi H. & Tutshini S. 2010. Hearing assessment, reliability, accuracy, and efficiency of automated audiometry. *Telemed J E Health*, 16(5), 557–563.

Wilber L.A., Kruger B. & Killion M.C. 1988. Reference thresholds for the ER-3A insert earphone. *J Acoust Soc Am,* 83, 669–676.