International Journal of Audiology 2011; 50: 440–447

Original Article

AMTAS®: Automated method for testing auditory sensitivity: III. Sensorineural hearing loss and air-bone gaps

Robert H. Margolis* & Brian C. J. Moore†

*Department of Otolaryngology, University of Minnesota, Minneapolis, USA; and Audiology Incorporated, Arden Hills, Minnesota, USA,
†Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge, UK

Abstract

Objective: The objectives were to measure the occlusion effect produced by three earphones—circumaural, supra-aural, and insert—and to compare air- and bone-conduction thresholds obtained with manual and automated methods for subjects with sensorineural hearing loss. *Design*: Acoustic and psychoacoustic occlusion effects were measured with each earphone. Manual and automated, air- and bone-conduction thresholds were compared. *Study sample*: Occlusion effects were measured for six adult subjects with normal external and middle ears. Pure-tone thresholds were measured for nineteen ears of thirteen subjects with sensorineural hearing loss. *Results*: The supra-aural earphone produced the largest occlusion effects, followed by the insert and circumaural earphones. Some systematic differences in air-conduction thresholds were found for the two procedures that may be attributable to earphone differences. A large air-bone gap at 4 kHz, reported in a previous study, was replicated. *Conclusions*: From 0.5 to 8.0 kHz, occlusion effects produced by the circumaural earphone are sufficiently small that covering the ear does not appreciably alter bone-conduction thresholds. Air-conduction threshold differences warrant further study to determine if reference equivalent threshold sound pressure levels for the two earphones produce equivalent thresholds. The large air-bone gap at 4 kHz suggests the possibility of an incorrect reference equivalent threshold force level at that frequency.

Sumario

Objetivo: Los objetivos fueron medir el efecto de oclusión producido por tres auriculares – circumaural, supra-aural y de inserción – y comparar umbrales de conducción aérea y ósea obtenidos con métodos manuales y automatizados para sujetos con hipoacusia sensorineural. *Diseño:* Se midieron los efectos de oclusión acústicos y psicoacústicos con cada auricular. Se compararon los umbrales de conducción aérea y ósea manuales y automatizados. *Muestra del Estudio:* Se midieron los efectos de oclusión para seis sujetos adultos con oídos externos y medios normales. *Resultados:* Los auriculares supra-aurales produjeron los mayores efectos de oclusión, seguidos de los auriculares de inserción y los circumaurales. Se encontraron algunas diferencias sistemáticas en los umbrales de conducción aérea para los dos procedimientos que pueden atribuirse a las diferencias en los auriculares. Una gran brecha aéreo-ósea en 4 kHz que se reportó en un estudio previo, fue replicada. *Conclusiones:* De 0.5 a 8 kHz, los efectos de oclusión producidos por los auriculares circumaurales son lo suficientemente pequeños por lo que cubrir el cubrir el oído no altera apreciablemente los umbrales de conducción ósea. Las diferencias en los umbrales de conducción aérea exigen estudios adicionales para determinar si los niveles umbrales equivalentes de referencia de presión sonora para los dos auriculares producen umbrales equivalentes. La mayor brecha aéreo-ósea en 4 kHz sugiere la posibilidad de un nivel de referencia equivalente de fuerza umbral en esa frecuencia.

Key Words: Audiometry; Automated audiometry; Hearing; Hearing test; Air conduction; Bone conduction; Threshold

Although automated pure-tone audiometry has been discussed in the audiology literature for decades (e.g. Rudmose, 1964), it has not been used widely in diagnostic hearing assessment. Recently there has been renewed interest due to the development of new methods and hardware (Ho et al, 2009; Margolis et al 2007, 2010, 2011; Swanepoel et al 2010; Swanepoel & Biagio, 2011). In general, these studies indicate that carefully designed automated methods are capable of providing results that are in good agreement with those obtained by experienced audiologists.

In previous reports we compared air-conduction and bone-conduction thresholds for normal and hearing-impaired subjects obtained by conventional, manual clinical audiometry, and by AMTAS®, an automated method for obtaining an audiogram, including air-conduction and bone-conduction thresholds with contralateral masking (Margolis et al, 2007, 2010, 2011). Differences between manual and AMTAS thresholds were compared to differences in manual thresholds obtained by two different audiologists. The rationale for this comparison is that, ideally, differences between manual and automated testing should be similar to those obtained using repeat testing by two skilled audiologists. The results showed that distributions of AMTAS-manual differences were generally similar to those between audiograms obtained by two audiologists.

In a subsequent study (Margolis et al, 2011) we compared thresholds measured using AMTAS and manual air-conduction for 4–8 year old children and derived a predictive formula for estimating the

---

Correspondence: Robert H. Margolis, University of Minnesota, Department of Otolaryngology, MMC396, Minneapolis, MN 55455, USA. E-mail: margo001@umn.edu

(Received 3 September 2010; accepted 21 March 2011)

ISSN 1499-2027 print/ISSN 1708-8186 online
DOI: 10.3109/14992027.2011.575085

Abbreviations

AMTAS	Automated method for testing auditory sensitivity
ANOVA	Analysis of variance
dB	Decibel
HL	Hearing level
RETFL	Reference equivalent threshold force level
RETSPL	Reference equivalent threshold sound pressure level
s.d.	Standard deviation
SNHL	Sensorineural hearing loss

accuracy of air-conduction audiograms. The results indicated that when audiograms with poor predicted accuracy are removed, the agreement between AMTAS and manual audiograms from children is comparable to the agreement between air-conduction audiograms obtained by two audiologists testing the same adult subjects.

There are two important procedural requirements for efficient automated testing. First, it is desirable to use earphones that do not produce occlusion effects. The occlusion effect is a shift (usually an increase) in the level of a bone-conducted signal resulting from occlusion of the ear canal. It can be measured as a threshold shift (usually a lower threshold) or as an increase in the ear-canal sound pressure level measured medially to the occlusion. Because bone-conduction threshold norms are based on the unoccluded condition, the shifted threshold produced by the occlusion effect is regarded as an artifactually altered value. For this reason, it is standard procedure to measure bone-conduction thresholds with the test ear uncovered. If the earphone does not produce an occlusion effect, both ears can be covered with earphones (for delivery of masking noise) without affecting bone-conduction thresholds. This avoids the need to reposition the masker earphone when switching the test ear and allows the entire audiogram (air- and bone-conduction with contralateral masking) to be obtained without interruption. If it were necessary to reposition earphones during the test, this would remove one of the advantages of automated testing.

The most commonly used audiometric earphone (Telephonics TDH type) is a supra-aural earphone that produces significant occlusion effects at 1 kHz and below (Dean & Martin, 2000; Stenfelt & Reinfeldt, 2007). Insert earphones used for audiometry, most commonly the Etymotic Research ER3A, also cause significant occlusion effects unless they are inserted deeply into the osseous portion of the ear canal (Dean & Martin, 2000; Stenfelt & Reinfeldt, 2007). Deep insertion for routine testing is impractical for reasons of comfort and safety. If the earphone is built into a circumaural muff that captures a large volume of air, the occlusion effect can be avoided (Stenfelt & Reinfeldt, 2007). For this reason, the Sennheiser HDA 200 circumaural earphone was selected for use in AMTAS testing.

The second important procedural consideration for automated testing is the placement of the bone vibrator. Although mastoid placement is commonly used for manual audiometry, forehead placement is preferred for automated testing to avoid the need for repositioning the bone vibrator during the test. There are several other advantages of forehead placement, as described by Margolis et al (2010).

This investigation was undertaken to address three specific issues related to the comparison of AMTAS and manual audiometry. First, we wished to confirm that occlusion effects are negligible for the Sennheiser HDA 200 earphones. Occlusion effect measurements were obtained with three commonly used audiometric earphones, including the HDA 200. Although Stenfelt and Reinfeldt (2007) reported model predictions of the occlusion effect for circumaural earphones similar to the HDA 200 earphone, we are not aware of any published measured data on the occlusion effect for this earphone.

The second issue addressed by this paper is whether audiometric thresholds are influenced by the type of earphone used. Until recently circumaural earphones such as the HDA 200 were not commonly used for manual audiometry (except for extended high-frequency testing). As noted earlier, Margolis et al (2010) reported similar air-conduction thresholds for AMTAS and for manual audiometry when both were obtained using the HDA 200 earphones. However, it is possible that thresholds would differ for AMTAS and manual audiometry if the latter were performed using the more commonly used supra-aural earphones. That possibility was assessed here. We reasoned that any differences between thresholds obtained using AMTAS and HDA 200 earphones on the one hand, and manual audiometry and TDH earphones on the other hand, could probably be attributed to earphone differences rather than procedure, bearing in mind the finding of Margolis et al (2010) of similar air-conduction thresholds for AMTAS and for manual audiometry when both were obtained using the HDA 200 earphones.

The third issue addressed by this paper was the air-bone gaps at 4 kHz for people with sensorineural hearing loss that have been reported in the literature for almost three decades (Frank & Holmes, 1981; Lightfoot & Hughes, 1993; Margolis et al, 2010) and are commonly discussed by audiologists (see for example http://www.aud.org.uk/forum/showthread.php?t = 946). We sought to replicate these air-bone gaps, to further quantify their magnitude, and to consider whether a change is needed in the 4-kHz reference equivalent threshold force levels (RETFLs) in the international (ISO 389-8, 2004) and American (ANSI S3.6-2010) standards for audiometers. This is an important issue because erroneous air-bone gaps at 4 kHz can lead to inappropriate follow-up and treatment for conductive hearing loss.

Study 1: Occlusion effects

Subjects

Six adult subjects ($>$21 years of age) participated in this study, three male and three female. All had normal hearing (thresholds $\leq$20 dB HL at 0.5, 1.0, and 2.0 kHz), normal otoscopic exams, and no recent history of middle-ear disease. The right ear of each subject was tested.

Methods

Occlusion effects were measured with the three earphone types commonly used for audiometry: the Telephonics TDH50 supra-aural earphone (Type 51 cushion); the Sennheiser HDA 200 circumaural earphone; and the Etymotic Research ER3A insert earphone. The ER3A earphone was coupled to the ear with one of two foam tips with the following dimensions when in the uncompressed condition: adult size: length = 14.0 mm; diameter = 13.0 mm; pediatric size: length = 14.0 mm; diameter = 9.0 mm. Two methods were used to measure the occlusion effect: an *acoustic method*, and a *psychoacoustic method*.

Acoustic occlusion effects were measured by recording the ear-canal sound pressure produced by a bone vibrator (Radioear B-71) placed on the forehead with the ear canal open and with the ear canal occluded by each of the earphones. The bone vibrator was coupled to the forehead with an elastic band designed to produce the appropriate force for audiometric measurements (5.4 ± 0.5 N), as specified in the audiometer standards (ISO 389.3, 1994; ANSI S3.6-2010). Ear-canal sound pressure was measured with a probe microphone with the probe tube placed in the ear canal to a depth of approximately one half to three quarters of the length of the canal. The output of the probe microphone was measured by a test system designed for electroacoustic measurements of hearing aids (Audioscan Verifit VF-1 with software version 3.0, and associated probe microphone).

Ear-canal sound pressure was initially recorded using a sweep-frequency tone with the ear canal open. Then, the earphone was placed over or in the ear without disturbing the probe tube and the measurement was repeated. For the ER3A insert earphone, measurements were made with two insertion depths of the foam insert, referred to as full and partial insertion. For the full insertion, the foam insert was placed so that the lateral edge was flush with the ear canal opening. For the partial insertion, approximately one-half of the length of the foam insert was beyond the ear canal opening. The differences between the sound pressures measured in the occluded and unoccluded conditions are reported for audiometric test frequencies from 0.25 to 6.0 kHz.

Psychoacoustic occlusion effects were measured for octave audiometric frequencies from 0.25 to 4.0 kHz. Audiometric thresholds were measured manually with a Madsen Aurical clinical audiometer, using bone-conducted signals delivered to the forehead by a Radioear B71 bone vibrator. Thresholds were measured first with the test ear open and then with the earphone on or in the ear. The insert earphone foam insert was placed with full insertion, as described above. The other earphone of the pair delivered masking noise at 30 dB effective masking level to the non-test ear to ensure that the test signals were audible only in the test ear. The difference between the two threshold measurements was taken as the psychoacoustic occlusion effect.

Results and Discussion

Means and standard deviations (s.d.) of acoustic and psychoacoustic occlusion effect measurements are provided in Table 1 and Figure 1. As previously reported, the occlusion effect is mostly a low-frequency phenomenon, with minimal effects above 1.0 kHz. In general acoustic occlusion effects were largest for the supra-aural earphone (TDH50), followed by the insert earphone (ER3A) with partial insertion, and the insert earphone with full insertion. The effects were smallest for the circumaural earphone (HDA 200). Psychoacoustic occlusion effects were largest for the supra-aural earphone and similar for the insert earphone with deep insertion and the circumaural earphone. Acoustic occlusion effects tended to be slightly larger than psychoacoustic effects, presumably because psychoacoustic bone-conduction thresholds are not determined solely by the ear-canal sound pressure.

Stenfelt and Reinfeldt (2007) compared acoustic and psychoacoustic occlusion effects produced by insert and circumaural earphones with predictions of their analog model of the ear canal. The circumaural device used for occlusion effect measurements had a volume of 30 cm^3. In addition, model predictions were presented for a 200 cm^3 circumaural enclosure, which approximates the 201 cm^3 volume of the Sennheiser HDA 200 earphones used in this study (Berg O., personal communication). Their results indicate that (1) acoustic occlusion effects are larger than psychoacoustic occlusion effects at low frequencies ($<$0.3 kHz), and (2) the magnitude of the occlusion effect is inversely related to the volume of a circumaural enclosure. Elpern and Naunton (1963) also reported an inverse relationship between the occlusion effect and enclosure volume.

Figure 2 shows the predictions of Stenfelt and Reinfeldt along with our acoustic occlusion effect measurements for the HDA 200. The model predictions show a 4-dB occlusion effect at 0.25 kHz, small negative occlusion effects in the range 0.5–2.0 kHz, and negligible effects at higher frequencies. The measured values show a 10-dB effect at 0.25 kHz but are on average within 2 dB of the predicted values, ranging from –1.8 to 1.3 dB, over the frequency range 0.5–4.0 kHz.

These results suggest that, for frequencies of 0.5 kHz and above, occlusion effects produced by the HDA 200 circumaural earphone are sufficiently small that thresholds obtained with the ears covered can be considered to be equivalent to those obtained with the ears open. The ER3A insert earphone, when partially inserted, produces threshold shifts of about 10 dB at 1 kHz and below, large enough to

Table 1. Acoustic and psychoacoustic occlusion effects for six adult subjects. Occlusion effects were measured for insert earphones (Etymotic Research ER3) with full insertion (F) and partial insertion (P), circumaural earphones (Sennheiser HDA 200), and supra-aural earphones (Telephonics TDH-50 with Type 51 cushion).

Frequency (kHz)		0.25	0.5	0.75	1.0	1.5	2.0	3.0	4.0	6.0
Earphone/Acoustic										
ER3 (F)	Mean	5.3	5.2	0.2	2.8	−0.5	−8.7	−2.5	−0.5	−0.8
	s.d.	4.1	3.3	4.5	4.4	2.5	2.7	4.4	2.4	1.7
ER3 (P)	Mean	10.5	10.3	7.2	9.8	4.2	−5.5	−2.7	−1.0	−0.2
	s.d.	6.1	3.0	6.0	5.6	3.1	3.2	3.1	2.2	0.8
HDA200	Mean	9.7	1.3	0.0	−1.8	−1.7	−0.5	0.3	0.3	1.0
	s.d.	4.3	6.1	4.3	4.4	3.4	1.9	2.1	2.7	1.3
TDH50	Mean	18.2	13.0	3.2	5.0	1.5	−2.2	1.5	3.0	2.5
	s.d.	8.2	6.1	8.3	10.4	6.6	8.0	5.0	5.9	6.9
Earphone/Psychoacoustic										
ER3 (F)	Mean	3.3	4.2	2.5	−0.8		−1.7		0.0	
	s.d.	6.1	7.4	5.2	5.8		4.1		0.0	
HDA200	Mean	5.0	1.7	0.8	−0.8		0.8		0.0	
	s.d.	4.5	6.8	8.0	7.4		3.8		3.2	
TDH50	Mean	16.0	15.0	8.0	3.0		−1.0		1.0	
	s.d.	4.2	3.5	2.7	4.5		2.2		2.2	

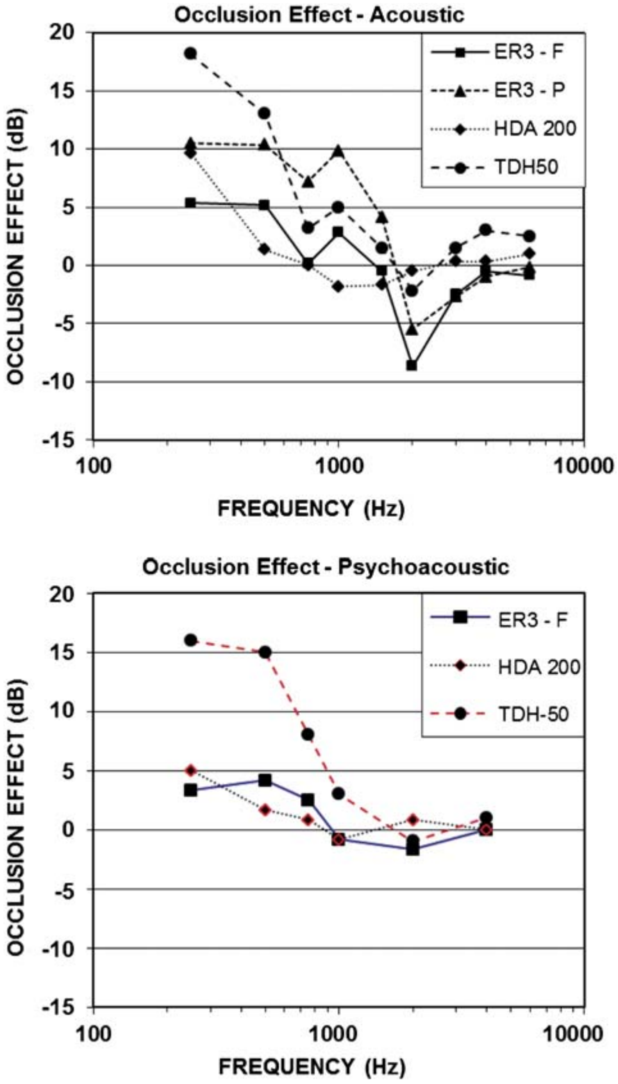

Figure 1. Mean acoustic and psychoacoustic occlusion effects for six adult subjects.

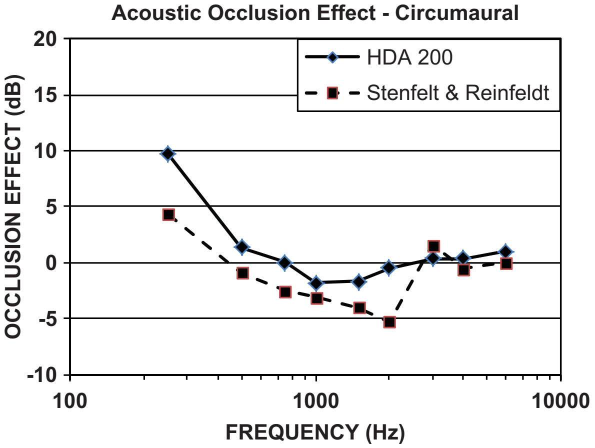

Figure 2. Mean acoustic occlusion effects for six subjects for the Sennheiser HDA 200 earphone and predicted occlusion effects for a 200 cm^3 circumaural enclosure from Stenfelt & Reinfeldt (2007).

cause significant errors in bone-conduction threshold measurements. The magnitude of the occlusion effect is critically dependent on the insertion depth and the type of coupling tip that is used (Stenfelt & Reinfeldt, 2007). A deeper insertion produces smaller occlusion effects at low frequencies than a more shallow insertion, but not as small as obtained with circumaural earphones. Negative acoustic occlusion effects were observed at 2 kHz with a deep insertion. There was no negative effect in the psychoacoustic measurements, presumably because the main route of sound to the cochlea at this frequency was not via the ear canal.

Study 2: Comparison of AMTAS and manual thresholds

Methods

AMTAS uses a single interval, Yes-No psychophysical bracketing procedure with feedback. Threshold is defined as the higher of two levels within a 5-dB interval at which a ‘Yes’ response follows a ‘No’ response. Thresholds are retested when certain quality indicators indicate that the initial threshold determination may be in error. Catch trials are presented randomly throughout the test to allow a quantitative measurement of false-alarm rate. Masking noise is always presented to the non-test ear by a proprietary method that estimates the appropriate masker level from the signal level and interaural attenuation of the transducers. Interaural attenuation values for the HDA 200 earphone were measured in our laboratory and are very similar to those reported recently by Brännström and Lantz (2010). At the conclusion of the test, AMTAS identifies ‘masking alerts’, thresholds for which overmasking or undermasking may have occurred. From these and other quality indicators, a predicted accuracy is calculated (Margolis et al, 2007).

Manual audiometry was performed by a research audiologist with many years of hearing assessment experience. She was instructed to use the Hughson-Westlake method, as she normally does during a clinical hearing evaluation. The threshold definition in the American standard for manual pure-tone audiometry was employed: ‘Threshold is defined as the lowest hearing level at which responses occur in at least one-half of a series of ascending trials, with a minimum of two responses out of three required at a single level’ (ANSI 3.21-2004). A plateau masking method was used for both air-conduction and bone-conduction testing.

AMTAS thresholds were obtained with a PC-based audiometer (Madsen Aurical) with circumaural earphones (Sennheiser HDA 200), and a standard clinical bone vibrator (Radioear B71) secured on the forehead with an elastic band designed to exert the proper force. Air-conduction stimuli were calibrated to international (ISO 389-8, 2004) and American (ANSI S3.6- 2010, Annex C) specifications with the earphone placed on an ear simulator with a flat-plate adapter (IEC 60318-1, 1998). Manual audiometry was performed by the research audiologist with a clinical audiometer (Grason Stadler GSI 61) and supra-aural earphones (Telephonics TDH-50 with Type 51 cushions) calibrated on an ear simulator (IEC 60318-3, 1998) according to international (ISO 389-1, 1998, Table 2) and American (ANSI S3.6-2010, Table 6) standards. Bone-conduction thresholds were measured with the bone vibrator secured to the mastoid with the spring band supplied with the audiometer. For both AMTAS and manual audiometry, the bone vibrator was calibrated to standard RETFLs. The vibrator was coupled to an artificial mastoid simulator (Larson-Davis AMC493) and calibrated to forehead RETFLs

Table 2. Average absolute differences: Audiologist vs. Audiologist, and AMTAS vs. manual.

	Two audiologists: 95% confidence interval	*Cambridge*	*Current study*
Air conduction	2.3–6.0	3.6	6.6
Bone conduction	2.9–7.9	7.7	7.7

(AMTAS) or mastoid RETFLs (manual audiometry) as specified in the international (ISO 389-3, 1994) and American (ANSI S3.6-2010) audiometer standards for bone-conduction stimuli. All testing was performed in a two-room, double-wall sound attenuating chamber that met standard requirements for permissible noise levels for audiometric testing (ANSI S3.1-1999).

Previous studies comparing automated and manual audiometry have based the comparison on two measures of the differences in thresholds obtained by the two methods (Margolis et al, 2007, 2010, 2011; Swanepoel et al, 2010). The average difference is the mean of the differences between pairs of thresholds for each frequency (AMTAS threshold minus manual threshold) and it gives an indication of systematic differences that are not explained by inherent variability in the measurements. The average absolute difference is the average of the absolute values of the differences (without regard to sign) between threshold pairs for the entire audiogram. It is typically larger than the average difference because it includes the effect of inherent variability in the threshold measures. Both average differences and average absolute differences are important in understanding the relationship between thresholds obtained with the two methods.

Subjects

Subjects were recruited from a pool of previously-tested research subjects and from clinic patients. The hearing loss was considered to be sensorineural if the average air-conduction, high-frequency pure-tone threshold average (2.0, 4.0, 8.0 kHz) was at least 35 dB HL (mean = 53 dB HL; range = 35–70 dB HL), and the air-bone gap, as measured using manual audiometry, did not exceed 5 dB at two or three frequencies (0.5, 1.0, 2.0 kHz), and did not exceed 10 dB at any of these three frequencies. Nineteen ears of thirteen subjects met these criteria. All subjects were adults (21–65 years of age).

Results and Discussion

The mean air-conduction audiograms for the two earphones/methods are shown in Figure 3. Although the trend in the average thresholds was similar for AMTAS and for manual audiometry, AMTAS thresholds were higher than manual thresholds by about 7 dB at 0.25, 0.5, 1.0, and 2.0 kHz, with smaller differences at higher frequencies. An analysis of variance (ANOVA) showed that the effect of earphones/method was statistically significant ($F(1,5) = 8.04$; $p = 0.005$). The effect of frequency was also significant ($F(1,5) = 67.2$, $p < .001$). The interaction between frequency and earphones/method was not significant ($F(1,5) = 0.46$; $p = 0.80$).

Average differences and average absolute differences between AMTAS and manual air-conduction thresholds are shown in Figure 4, which shows the results of Margolis et al (2010) for comparison (Cambridge). Average differences for the present study (top panel) correspond to the differences in the average air-conduction audiograms shown in Figure 3. The close agreement between AMTAS and manual air-conduction thresholds obtained when both procedures were performed with the same earphone (Cambridge) leads us to hypothesize that the differences in the present study are due to earphone differences. This may indicate that the reference equivalent threshold sound pressure levels (RETSPLs) in the international and American audiometer standards need adjustment to obtain equivalence between the two earphones. The study should be replicated with a larger number of subjects to determine if a change in the standards should be made.

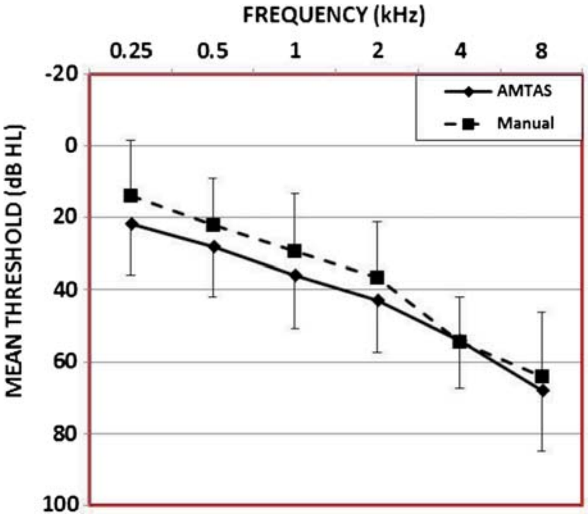

Figure 3. Mean air-conduction thresholds for 19 ears of 13 subjects with sensorineural hearing loss. Thresholds were obtained by AMTAS and by manual audiometry. Error bars are 1 standard deviation.

The average absolute differences between AMTAS and manual air-conduction thresholds were only slightly larger than the average differences (compare diamonds in the top and bottom panels of Figure 4). This suggests that the absolute differences are largely accounted for by the systematic differences evident in Figure 3.

Average differences and average absolute differences between bone-conduction thresholds measured using AMTAS and manual audiometry are shown in Figure 5, which shows the results of Margolis et al (2010) for comparison (Cambridge). Average differences (top panel) across the four test frequencies were about 5 dB, indicating that thresholds obtained with AMTAS were higher than those obtained using manual audiometry. In contrast, small negative differences were obtained in the Cambridge study. These differences were unexpected because the bone-conduction procedures were identical in the two studies. It is possible that differences in the calibration procedure for bone-conduction stimuli in the two studies contributed to these differences. The bone-conduction stimuli in the Cambridge study were calibrated using a Bruel and Kjær artificial mastoid (type 4930); the stimuli in the current study were calibrated using a Larson-Davis artificial mastoid simulator (AMC493). Although the devices should yield equivalent results, the use of two different devices may have contributed to the difference in the results. Average absolute differences (bottom panel of Figure 5) were in better agreement across studies, with differences ranging from 5 to 8 dB across the four test frequencies.

In previous reports (Margolis et al, 2007, 2010, 2011) it has been suggested that the validity of AMTAS thresholds can be assessed by

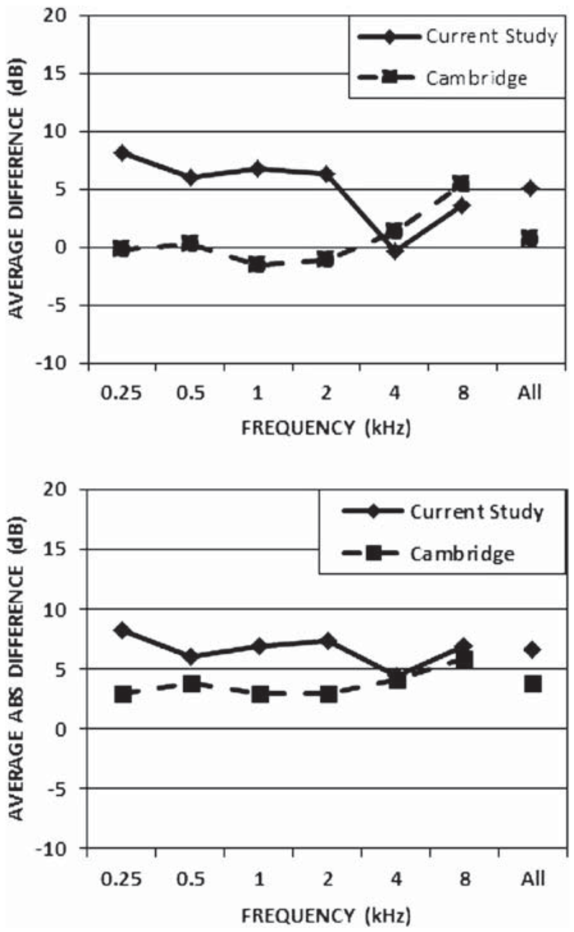

Figure 4. Average differences (top panel) and average absolute differences (bottom panel) between AMTAS and manual air-conduction thresholds from the current study and from Margolis et al (2010, Cambridge).

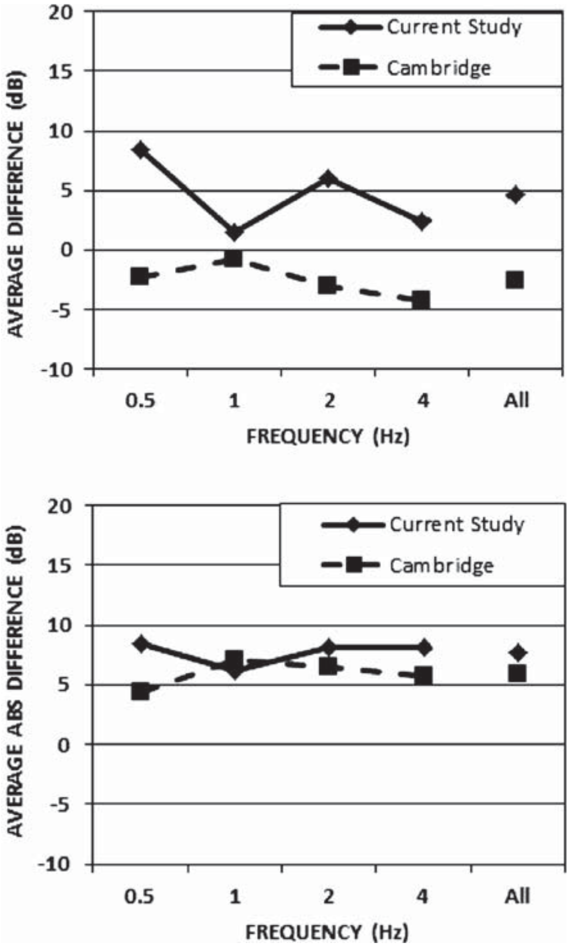

Figure 5. Average differences (top panel) and average absolute differences (bottom panel) between AMTAS and manual bone-conduction thresholds from the current study and from Margolis et al (2010, Cambridge).

comparing the average absolute differences between AMTAS and manually obtained thresholds to the differences obtained by two expert audiologists testing the same patients. If the average absolute differences for AMTAS and manual thresholds are within a confidence interval determined from differences for two audiologists, this would suggest that AMTAS performs similarly to an audiologist.

Table 2 shows average absolute differences for two audiologists, and for AMTAS-manual threshold differences from the Cambridge study and the current study. Results from the Cambridge study are well within the 95% confidence interval for two audiologists for air conduction, and just within the interval for bone conduction. Average absolute differences for the current study fall slightly outside the confidence interval for air conduction and just within it for bone conduction. The previously mentioned earphone difference is probably responsible for the discrepancy for air-conduction.

AIR-BONE GAPS

In the Cambridge study (Margolis et al, 2010) there were significant air-bone gaps at 4 kHz for subjects with sensorineural hearing loss. It was suggested that the standard RETFLs for both mastoid and forehead placement may need adjustment to avoid results suggesting conductive hearing loss when there was, in fact, no such loss. The results of the present study replicate that finding. Figure 6 shows air-bone gaps from the current study and from the Cambridge study for four test frequencies, obtained with AMTAS and with manual methods. Also shown are the means of the four determinations, weighted by the number of subjects contributing to each average (circles).

Average air-bone gaps are expected to be zero for patients with sensorineural hearing loss and completely normal middle-ear function. However, the distribution of air-bone gaps is determined by the combined variability of air-conduction and bone-conduction thresholds. Studebaker (1967) assumed that the air-bone gap is a normally distributed variable with a standard deviation of 5 dB. On this basis, he predicted that individual air-bone gaps would be 0 dB only 38% of the time, given the usual 5-dB step size of clinical measurements. Margolis (2008) suggested, based on published measures of the variability of air- and bone-conduction thresholds, that the standard deviation of the air-bone gap is about 8.6 dB. In this case, the air-bone gap would be 0 dB only 21% of the time, and air-bone gaps of 10 dB or more (both positive and negative) would occur 40% of the time. Despite this large expected range of air-bone gaps, the average measured air-bone gap at 4 kHz, as shown in Figure 6, is well outside the range that would be expected from random variability.

Dirks et al (1979) obtained the normative data from which the bone-conduction RETFLs were derived. Using those norms, data

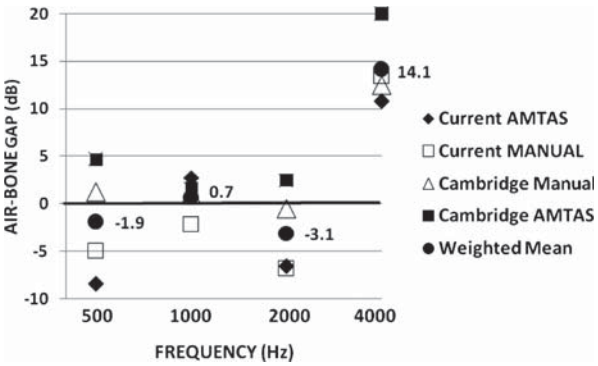

Figure 6. Mean air-bone gaps for AMTAS and manual thresholds obtained in the Cambridge study and the current study. Circles are weighted means of the four values at each frequency.

from patients with sensorineural hearing loss showed near-zero air-bone gaps at all test frequencies (0.25–8.0 kHz). At 0.5, 1.0, and 2.0 kHz, the small deviations from zero in Figure 6 are close to the results of Dirks et al. The differences can plausibly be accounted for by the inherent variability of threshold measurements (Studebaker, 1967; Margolis, 2008). At 4.0 kHz, however, the average air-bone gap of 14.1 dB differs markedly from the data of Dirks et al, and is not expected from the variability discussed by Studebaker (1967) and Margolis (2008). Rather, it suggests that an adjustment in the RETFL may be required to eliminate erroneous air-bone gaps at that frequency. The good agreement between 4-kHz air-conduction thresholds obtained with the two earphones (Figure 4, top panel) suggests that the problem is in the bone-conduction RETFL and not the air-conduction RETSPL. We propose that an adjustment of about $-$ 14 dB in the 4-kHz RETFL should be considered to eliminate the air-bone gaps seen at this frequency in listeners with sensorineural hearing loss. The RETFL at 4.0 kHz specified in the international (ISO 389-3, 1994) and American (ANSI S3.6-2010) standards is 35.5 dB re 1 μN for mastoid placement and 43.5 dB re 1 μN for forehead placement. An adjustment of -14 dB would bring these values to 21.5 and 29.5 dB re 1 μN for forehead and mastoid placement, respectively. However, the exact value of the required adjustment needs to be determined in a more extensive study designed for that purpose.

Summary and Recommendations

In Study 1, occlusion effects were measured using two methods (acoustic and psychoacoustic) with three earphones (Sennheiser HDA 200 circumaural, Telephonics TDH50 supra-aural, and Etymotic Research ER3A insert). The supra-aural earphone produced the largest occlusion effects, followed by the insert earphone with shallow insertion, the insert earphone with deep insertion, and the circumaural earphone. The Sennheiser HDA 200 circumaural earphone produced negligible occlusion effects over the range 0.5 to 4.0 kHz.

In Study 2, AMTAS thresholds obtained using HDA 200 circumaural earphones were compared with manual thresholds obtained using TDH50 supra-aural earphones for subjects with sensorineural hearing loss. Air-conduction thresholds were systematically higher for the AMTAS/circumaural measurements than for the manual/supra-aural measurements by about 7 dB over the range 0.25–2.0 kHz, but thresholds did not differ markedly for higher frequencies.

Bone-conduction thresholds were obtained with mastoid placement for manual testing and forehead placement for AMTAS testing. Bone-conduction thresholds were, on average, 5 dB higher for AMTAS than for manual testing, a finding different from that of our previous study (Margolis et al, 2010), for which AMTAS thresholds were, on average, 3-dB lower than manual thresholds.

Air-bone gaps averaged over the two studies showed that small (negligible) air-bone gaps occurred at 0.5, 1.0, and 2.0 kHz, but the average air-bone gap at 4 kHz was 14.1 dB, suggesting a problem with the standard RETFL at that frequency.

These results support the following recommendations.

1. Circumaural earphones (specifically the Sennheiser HDA 200) produce minimal occlusion effects for frequencies of 0.5 kHz and above, a characteristic that is advantageous for bone-conduction testing because testing can be performed with the earphones covering the ears, eliminating the need to move the transducers during the test. With forehead bone-conduction transducer placement, the entire pure-tone audiogram (air conduction and bone conduction with contralateral masking) can be performed without moving the transducers.
2. The higher thresholds obtained with circumaural earphones than with supra-aural earphones suggest that an adjustment of the standard RETSPLs for one or both of the earphones is required for equivalence. This result should be replicated before an adjustment of the RETSPLs is considered. The replication should be consistent with the guidelines for determining reference threshold levels that are provided in ISO 389-9, 2009.
3. The large air-bone gap at 4.0 kHz, previously reported by Margolis et al (2010) and replicated here, suggests that the RETFL at that frequency may require adjustment. This should be explored further in a study that conforms to the ISO 389-9 guidelines.
4. This study and its predecessors (Margolis et al, 2010, 2011) provide three different comparisons of AMTAS and manual thresholds. In the study of Margolis et al (2010), circumaural earphones were used for both automated and manual testing. In the study of Margolis et al (2011, adult subjects) and the current study, circumaural earphones were used for AMTAS testing and supra-aural earphones were used for manual testing. A study that provides both comparisons in the same group of patients with normal hearing and various hearing loss configurations would clarify the origin of some of the differences observed between the two studies.

Acknowledgements

This research was supported by grant R42 DC007773 from the National Institutes of Health. We are grateful to research audiologist Allison Kohtz for collecting the data presented in this report, and to three anonymous reviewers for their helpful suggestions. AMTAS® intellectual property is owned by the first author and Audiology Incorporated, and may become a commercial product. Author BM was supported by the MRC (UK).

Declaration of interest: The first author declares a conflict of interest that is disclosed in the Acknowledgements section. The second author reports no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References

ANSI S3.21-2004 (R2009). 2009. Methods for manual pure-tone audiometry. American National Standards Institute, New York.

ANSI S3.1-1999. 1999. Maximum permissible ambient noise levels for audiometric test rooms. American National Standards Institute, New York.

ANSI S3.6-2010. 2010. Specification for audiometers. American National Standards Institute, New York.

Brännström K.J. & Lantz J. 2010. Interaural attenuation for the Sennheiser HDA 200 circumaural earphones. *Int J Audiol*, 49, 467–471.

Dean M.S. & Martin F.N. 2000. Insert earphone depth and the occlusion effect. *Am J Audiol*, 9, 131–134.

Dirks D.D., Lybarger S.F., Olsen W.O. & Billings B.L. 1979. Bone conduction calibration: Current status. *J Speech Hearing Dis*, 44, 143–155.

Elpern B. & Naunton R. 1963. The stability of the occlusion effect. *Arch Otolaryng*, 77, 376–384.

Frank T. & Holmes A. 1981. Acoustic radiation from bone vibrators. *Ear Hear*, 2, 59–63.

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

IEC 60318-1. 1998. Electroacoustics – Simulators of human head and ear – Part 1: Ear simulator for the calibration of supra-aural earphones. Geneva: International Electrotechnical Commission.

IEC 60318-3. 1998. Electroacoustics – Simulators of human head and ear – Part 3: Acoustic coupler for the calibration of supra-aural earphones used in audiometry. Geneva: International Electrotechnical Commission.

ISO 389.1. 1998. Acoustics – Reference zero for the calibration of audiometric equipment – Part 1: Reference equivalent threshold sound pressure levels for pure tones and supra-aural earphones. Geneva: International Organization for Standardization.

ISO 389.3. 1994. Acoustics – Reference zero for the calibration of audiometric equipment – Part 3: Reference equivalent threshold force levels for pure tones and bone vibrators. Geneva: International Organization for Standardization.

ISO 389-8. 2004. Acoustics – Reference zero for the calibration of audiometric equipment – Part 8: Reference equivalent threshold sound pressure levels for pure tones and circumaural earphones. Geneva: International Organization for Standardization.

ISO 389-9. 2009. Acoustics – Reference zero for the calibration of audiometric equipment – Part 9: Preferred test conditions for the determination of reference hearing threshold levels. Geneva: International Organization for Standardization.

Lightfoot G.R. & Hughes J.B. 1993. Bone conduction errors at high frequencies: Implications for clinical and medico-legal practice. *J Laryngol Otol*, 107, 305–308.

Margolis R.H. 2008. The vanishing air-bone gap: Audiology's dirty little secret. Audiology Online (www.audiologyonline.com/articles), October 20, 2008.

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., 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., Saly G.L., Le C. & Laurence J. 2007. Qualind™: A method for assessing the accuracy of automated tests. *J Amer Acad Aud*, 18, 78–89.

Rudmose W. 1964. Automated audiometry. In: J. Jerger, *Modern Developments in Audiology*. New York: Academic Press.

Stenfelt S. & Reinfeldt S. 2007. A model of the occlusion effect with bone-conducted stimulation. *Int J Audiol*, 46, 595–608.

Studebaker G.A. 1967. Intertest variability and the air-bone gap. *J Speech Hear Dis*, 32, 82–86.

Swanepoel de W. & Biagio L. 2011. Validity of diagnostic computer-based air and forehead bone conduction audiometry. *J. Occup Environ Hyg*, in press.

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