Wideband Reflectance in Normal Caucasian and Chinese School-Aged Children and in Children with Otitis Media with Effusion

Alison N. Beers,1 Navid Shahnaz,2 Brian D. Westerberg,3 and Frederick K. Kozak4,5

Objectives: Wideband reflectance (WBR) is a middle ear analysis technique that quantifies frequency-specific sound conduction over a wide range of frequencies. One shortcoming of WBR is that there is limited normative data, particularly for pediatric populations and children with middle ear pathology. The goals of this study were to establish normative WBR data for early school-aged children; to determine whether WBR differs significantly between Caucasian and Chinese children, male and female children, and children and adults (experiment 1); and to compare the normative pediatric WBR data with the WBR data obtained from children with abnormal middle ear conditions (experiment 2).

Design: WBR was measured from 78 children with normal middle ear status with an average age of 6.15 yrs and 64 children with abnormal middle ear status with an average age of 6.34 yrs. Control group subjects and subjects without previously diagnosed middle ear pathology were recruited from eight elementary schools in the Greater Vancouver Area. Subjects with known middle ear pathology were recruited through the British Columbia Children’s Hospital Otalaryngology department. Middle ear effusion (MEE) was identified in one of the two ways. In the British Columbia Children’s Hospital group, MEE was diagnosed by a pediatric otolaryngologist (OTL) using pneumatic otoscopy and video otomicroscopy. These cases (21 ears) were classified as OTL confirmed. Subjects who were assessed through screenings at their elementary schools and suspected to have MEE based on audiological test battery results including elevated air conduction thresholds, flat low- and high-frequency tympanograms, and absent transient-evoked otoacoustic emissions were classified as not OTL confirmed (21 ears). Data were statistically analyzed for effects of gender, ethnicity (Caucasian versus Chinese), age (child versus adult), and middle ear condition. WBR equipment used for this study was from Mimosa Acoustics (RMS-system, version 4.03). Data were averaged in one-third octave bands collected from 248 frequencies ranging from 211 to 6000 Hz.

Results: Control group subject data (experiment 1) revealed no effects of gender or ear, and their interactions with frequency were not significant. There was a significant interaction between ethnicity (Caucasian versus Chinese) and frequency. Chinese children had lower energy reflectance (ER) values over the mid-frequency range. ER was significantly different between pediatric data and previously collected adult data. Diseased group ER was significantly different among all four middle ear conditions (normal, mild negative middle ear pressure, severe negative middle ear pressure, and MEE) (experiment 2). The overall test performance of ER was objectively evaluated using receiver operating characteristic (ROC) curve analyses; it was compared across frequencies averaged in one-third octave bands. Statistical comparison of the area under ROC (AUROC) plots revealed that ER above 800 Hz (except for ER at 6300 Hz) had better test performance in distinguishing normal middle ear status from MEE compared with ER at 630 and 800 Hz. Although not statistically different from other frequencies between 800 and 5000 Hz, ER at 1250 Hz had the largest AUROC curve (sensitivity of 96% and specificity of 95%) and was selected for further analysis. Comparison of AUROC curves between WBR at 1250 Hz and static admittance at 226-Hz probe tone frequency revealed significantly better test performance for WBR in distinguishing between healthy ears and MEE.

Conclusions: A preliminary set of normative ER data have been generated for a pediatric population between the ages of 5 and 7 yrs, which were significantly different from previously gathered normative adult ER data. In this study, pediatric normative data were warranted for testing children, but ethnic-specific norms were not required to detect middle ear pathology and changes in middle ear status. WBR shows promise as a clinical diagnostic tool for measuring the mechaanoacoustic properties of the middle ear and the changes that result in the presence of negative middle ear pressure or MEE.

(End & Hearing 2010;31:221–233)

INTRODUCTION

Otitis media (OM) is the second most prevalent childhood disease after the common cold, with an estimated 5 million cases of OM in the United States per year responsible for most antibiotic prescriptions in children (Hendley 2002) at an annual cost of 4 billion dollars (U.S.) (Stool 1994). OM with effusion (OME), where fluid is present behind the eardrum without signs of active infection is thought to represent between 25 and 35% of the total cases of OM (Sagrades et al. 1992; Stool 1994). Eighty percent of children experience at least one episode of OME (Mora et al. 2002) with the highest occurrence in children younger than 6 yrs (Thrasher & Allen 2005). Chronic middle ear pathology and associated conductive hearing loss may disrupt normal speech, language, and cognitive development (Feldman & Gelman 1986; Nittouwer 1996; Mody et al. 1999; Shriberg et al. 2000; Cashy 2001). It has been shown that conductive hearing loss can reduce central auditory system resting neural activity in animals (Tucci et al. 2001) and can affect binaural processing in children (Hogan & Moore 2003).

The high prevalence of OME, the financial strain associated with its management, and the potential auditory and language delay that arise if left untreated have motivated the researchers to develop methods for its early and accurate diagnosis. Conventional (226 Hz) tympanometry is commonly used as a clinical predictor of middle ear effusion (MEE) in children (Nozza et al. 1992, 1994). Of all quantifiable parameters obtained from conventional 226-Hz tympanograms, tympanometric width (TW), or sharpness of the tympanogram shape, has been demonstrated to be the best single criterion for detecting MEE with sensitivity of 81% and specificity of 82% (Nozza et al. 1992, 1994). However, there is no consensus regarding which tympanometric parameter offers the best...
assessment of middle ear status. A study by Margolis et al. (1998), for example, found that TW was not an effective test for detecting significant middle ear pathology in simulated middle ear lesions in chinchillas. Conventional 226-Hz tympanometry is not able to detect sequelae and subtle changes in middle ear mechanics after OM (Hanks & Robinette 1993; Margolis et al. 1994; Vlachou et al. 2001). Harris et al. (2005) reported that of 21 cases of surgically confirmed MEE (via myringotomy), three were identified as normal by 226-Hz tympanometry.

Wideband reflectance (WBR) was developed to measure middle ear function over a wide range of frequencies (Keefe et al. 1992) and is now emerging as a useful clinical tool (Keefe & Levi, 1996; Margolis et al. 1999; Feeney et al. 2003; Allen et al. 2005; Vander Werff et al. 2007; Shahnaz 2008; Shahnaz et al. 2009). WBR is able to provide a more detailed assessment of the middle ear than conventional 226-Hz tympanometry, it does not require pressurization of the ear canal, it is able to predict the degree of conductive hearing loss, and the depth of probe insertion is not critical (Stinson et al. 1982; Voss & Allen 1994; Keefe & Levi 1996; Keefe et al. 1993, Piskorski et al. 1999; Huang et al. 2000; Feeney et al. 2003; Keefe & Simmons 2003). Energy reflectance (ER) is often measured in WBR research and represents the ratio of reflected to incident acoustic power presented via a probe inserted into the ear canal (Voss & Allen 1994). ER ranges from zero, indicating that all sound energy has been absorbed into the middle ear, to one, indicating that all sound energy has been reflected by the middle ear (Stinson 1990), and it is frequency dependant, when adult human ER is greatest below 1000 Hz and above 4000 Hz.

Evidence suggests a useful role for ER in the diagnosis of ears with OM (Keefe & Levi 1996; Feeney et al. 2003). Hunter et al. (2008) found improved test performance of ER in correctly identifying MEE in an infant population compared with conventional and high-frequency tympanometry, with a lower incidence of inconclusive results. A case study by Hunter and Margolis (1997) revealed that the presence of MEE resulted in abnormal ER when conventional tympanometry indicated normal middle ear admittance. Several studies have reported significantly higher ER in children and adults with OME compared with age-matched control subjects and that elevated ER occurs over a wide range of frequencies (Jeng et al. 1999; Feeney et al. 2003; Hunter et al. 2008). In addition to distinguishing between normal and pathological middle ears, WBR can provide information about the type of middle ear pathology (Allen et al. 2005). Unlike tympanometry, Piskorski et al. (1999) found that WBR is able to predict conductive hearing loss. The authors reported that ER between 2 and 4 kHz is a sensitive indicator of middle ear status and is a more accurate predictor of conductive impairment at 0.5 kHz than is ER at 0.5 kHz.

Comprehensive WBR studies to date involve primarily infant/toddler and adult subjects (Keefe et al. 2000; Feeney & Sanford 2005; Shahnaz & Bork 2006; Hunter et al. 2008; Shahnaz 2008). Previous research has identified that ethnicity differences in ER exist among Chinese and Caucasian adults (Shahnaz & Bork 2006), but it has not yet been determined whether these differences also exist among children of different ethnicities. The reliability and clinical effectiveness of WBR in detecting OME in school-aged children have not been extensively investigated. Data from a sufficiently large sample size are required before a representative normative data base can be established. The goal of this study was to explore the mechanoacoustical properties of the middle ear within a multiethnic pediatric population consisting of children with healthy middle ear status and those with varying degrees of middle ear pathology. Development of normative pediatric WBR data may enable this measurement technique to become a useful clinical diagnostic tool.

Three specific questions were addressed by this study:

1. What are the WBR patterns generated by a pediatric population with healthy middle ear status, and do they differ as a function of ear, gender, or ethnicity?
2. Do the WBR values measured from a pediatric population differ from the adult normative data obtained by Shahnaz and Bork (2006)?
3. How do the normative pediatric data compare with the WBR patterns of children with middle ear pathology, specifically MEE and negative middle ear pressure?

### MATERIALS AND METHODS

Institutional clinical research ethics board approval was obtained before commencement of the study, and informed consent was obtained from all subjects.

#### Experiment 1: Control Group Subjects

**Subjects** The control group consisted of 78 subjects (144 ears), 38 girls (69 ears) and 40 boys (75 ears) recruited from eight elementary schools in the Greater Vancouver Area. See Table 1 for the breakdown of ethnic origin within this group. Subjects ranged in age from 5 yrs 1 mo to 6 yrs 11 mo (average age, 6.15 yrs). Date of birth and ethnicity were reported by subjects’ parents. Permission was obtained from the participating schools and school boards.

**Inclusion Criteria for Control Group Subjects**

2. Pass a transient-evoked otoacoustic emission (TEOAE) screening. A pass consisted of a good emission-to-noise ratio of greater than or equal to 4.4 dB in both ears.

### TABLE 1. Summary of different ethnic distributions among middle ear conditions

<table>
<thead>
<tr>
<th>Middle ear condition</th>
<th>Total Ears</th>
<th>Caucasian</th>
<th>Chinese</th>
<th>Mixed or other ethnic origin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Controls</td>
<td>144</td>
<td>63</td>
<td>60</td>
<td>21</td>
</tr>
<tr>
<td>Mild negative middle ear pressure (−100 to −199 daPa)</td>
<td>30</td>
<td>15</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td>Severe negative middle ear pressure (−200 daPa or more negative)</td>
<td>24</td>
<td>17</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Middle ear effusion</td>
<td>42</td>
<td>28</td>
<td>6</td>
<td>8</td>
</tr>
</tbody>
</table>

[Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.]
ratio in four of five half-octave bands centered at 1, 1.5, 2, 3, and 4 kHz. The pass criterion was a signal-to-noise ratio of at least 3 dB at 1 and 1.5 kHz and 6 dB at 2, 3, and 4 kHz. This stopping criterion has been shown to have a reasonable test performance for detecting hearing loss in a large sample of newborns (Norton et al. 2000).

3. Normal tympanometric peak pressure (~99 daPa or less negative at 226 Hz using positive to negative pressure direction [Jerger 1970; Silman et al. 1992]).

4. No apparent pathology of either the outer or middle ear on otoscopic examination.

5. No evidence of a syndrome or developmental delay and no reported history of head trauma or hospitalization for an illness.

6. Student in either kindergarten or grade 1.

A modification of the ASHA (1997) guidelines was used for screening normal auditory function. Our criterion for normal auditory function was pure-tone thresholds of ≤20 dB HL from 500 to 4000 Hz. Screening results at 500 Hz are not a component of the ASHA guidelines for this age group but were included in this study to more effectively screen for conductive hearing loss. Testing in the elementary schools was conducted in the otherwise empty library or counseling office, which were quiet. Testing stopped during recess and lunch when hallway and playground noise could have confounded test results. Periodic listening checks (several times daily) revealed that all tones were audible at screening levels to normal-hearing testers. ASHA guidelines for 226-Hz tympanometry were used, where the passing criterion was static admittance of ≥0.3 mmho and TW ≤200 daPa.

Instrumentation • Pure-tone audiometry was conducted using a portable audiometer (Maico MA 40), which was calibrated according to American National Standards Institute (ANSI) standards (re: S3.6.1996). Supra-aural headphones (TDH-39) were used for all subjects.

TEOAEs were tested using an ILO-292 Analyzer (Ototronics, Ltd., Hatfield, England), which was calibrated based on the instructions provided in the operator’s manual from the manufacturer. Tympanometric measures were obtained using Grason Stadler Instrument (GSI) Tympstar Version 2 middle ear analyzer. The GSI system was calibrated using standard cavities according to the operation manual provided by the manufacturers and also in accordance with ANSI specifications.

WBR equipment used for this study was from Mimosa Acoustics (RMS-system, version 4.03). The WBR system and calibration procedure were similar to those described and used by Voss and Allen (1994) and Shahnaz and Bork (2006). Calibration of the WBR equipment was completed using the four-cavity calibration device of different sizes before the testing of each subject (Allen 1986). Measurement of ER depends on the cross-sectional area of the ear canal (S)* at the plane of measurement. The Mimosa acoustics system physically measures S based on the diameter of the probe tip size selected for measurement. During the initial step of calibration, the ER-10C Thevenin (Norton) parameters, source pressures, and impedance were measured in four cavities as was done by Allen (1986) and Voss and Allen (Jeng, Reference Note 1). The Thevenin equivalent parameters measured across the frequency range in each of the four cavities had to be at least 90% of the accepted range determined by the manufacturer or the probe tip was discarded. The frequency response measured during the calibration process was used to determine the transducer source pressure and source impedance in ear canal.

Procedure • All testing was conducted by a registered audiologist or by an audiology graduate student under the supervision of a registered audiologist. Otoscopic examination was performed to rule out any gross abnormalities of the external ear, ear canal, or tympanic membrane. A TEOAE screening was conducted to verify normal cochlear function at the level of the outer hair cells and as an indirect screening of middle ear status. Tympanometry was conducted with both conventional (226 Hz) and high-frequency (1000 Hz) probe tones, using a positive to negative pressure sweep direction. The static admittance and TW of the tympanograms were automatically calculated by the GSI system at 226 Hz using positive tail compensation. Aside from meeting ASHA (1997) guidelines for 226-Hz tympanometry, tympanograms were required to have a discernable peak at both 226- and 1000-Hz probe tone frequencies.

Behavioral audiological thresholds were screened at 20 dB HL at 1, 2, and 4 kHz according to the ASHA (1997) guidelines. Audiological thresholds were also screened at 20 dB HL at 500 Hz as middle ear pathology often results in an elevated air conduction threshold at this frequency. Normal-hearing status was ascertained in a particular ear after two clear hand raise responses (of a possible 3 responses) at each of the four test frequencies. Ambient noise was not measured before testing, but periodic listening checks conducted by the tester with normal hearing ensured that all of the tones were clearly audible.

The Mimosa Acoustics WBR system acquires ER data as a function of frequency and displays percentage values graphically. Data were collected at 248 frequencies from 211 to 6000 Hz and were separated into 23-Hz intervals. Only 14A foam WBR probe tips were used for this study. A calibration procedure was performed with each probe tip before use, and tips were inserted into the ear canal, so that they formed a seal with the canal walls.

A new software window opens for each ER measurement. This window shows the response spectrum and the noise levels as functions of frequency. The ER response can be remeasured until a smooth curve is obtained, and the levels of the signal are clearly higher than the levels of the noise at all frequencies (at least 3 dB above the noise floor at the lowest frequency measured). When an ER response is accepted, the measurement window closes and the results are graphically displayed on the four-window screen. Significant visual overlap of at least two measurement curves was required to ascertain a reliable representation of middle ear status.

To obtain numeric values for the ER data, the two overlap measures were imported into a Microsoft Excel spreadsheet, and one was selected randomly as the difference between the two was minimal (≤5% at low frequencies only).

*The ear-canal impedance normalized by its characteristic impedance $Z_c = \frac{\mu c S}{Z_o}$ where $Z_o$ is the characteristic impedance; $\mu$ is the density of air; $c$ is the speed of sound; and $S$ is the cross-sectional area of the ear canal (for more information see Voss & Allen, 1994).
Experiment 2: Diseased Group Subjects

Subjects • There were three subject groups with abnormal middle ear conditions:

1. Mild negative middle ear pressure group (−100 to −199 daPa);
2. Severe negative middle ear pressure group (−200 daPa or more negative); and
3. MEE group.

There were 21 subjects (30 ears) with mild negative middle ear pressure, 18 subjects (24 ears) with severe negative middle ear pressure, and 25 subjects (42 ears) with MEE. See Table 1 for a breakdown of the ethnic origin within the 3 diseased groups. Diseased group subjects ranged in age from 3 to 12 yrs; the mean age was 6.34 yrs. Date of birth and ethnicity were reported by subjects’ parents.

Subjects were recruited through two sources: diseased group subjects recruited through elementary schools were subjects with negative middle ear pressure, either mild or severe, or MEE. These subjects were initially recruited as control subjects but were found to have abnormal middle ear status at the time of testing. Subjects in this group were recruited and tested in the same manner as those in the control group. There were 21 ears in this group with MEE.

The second group of diseased group subjects was recruited through the Division of Pediatric Otolaryngology at British Columbia Children’s Hospital (BCCH). This group of subjects consisted only of ears with MEE (N = 21).

Instrumentation

Diseased Group Subjects Recruited through Elementary Schools • Subjects recruited through their elementary school were tested using identical instrumentation to that described above for control group subjects for screening of audiological thresholds and gathering TEOAE, tympanometric, and WBR data.

Diseased Group Subjects Recruited through BCCH • Audiological thresholds were measured using a GSIs 61 audiometer and calibrated bone oscillator for subjects tested through the BCCH Audiology Department. Testing was performed in a sound-treated booth. The GSI 61 audiometer and supra-aural headphones were calibrated according to ANSI standards (re: S3.6.1996). Daily biological listening checks were performed before testing. Instrumentation described in the control group section for gathering TEOAE, tympanometric, and WBR data were identical to instrumentation used to gather specified data from diseased group subjects tested at BCCH.

Procedure • Otoscopy was performed to rule out gross abnormalities of the external ear and tympanic membrane. TEOAE, tympanometric, and WBR data were gathered for all diseased group subjects using the same procedure outlined above for control group subjects. The static admittance and TW of the tympanograms were considered when drawing conclusions regarding middle ear condition; tympanograms were required to have a discernable peak with a TW ≤ 200 daPa if a value for negative middle ear pressure was to be assigned. Tympanograms with TW greater than 200 daPa were classified as flat. High-frequency (1000 Hz) tympanograms were evaluated qualitatively and were used to assist in the judgment of the 226-Hz tympanograms that were difficult to interpret.

Behavioral audiometry results were obtained in two ways, according to the test environment:

Diseased Group Subjects Recruited through Elementary Schools • Subjects tested at their elementary school had audiological thresholds screened at 20 dB HL at 0.5, 1, 2, and 4 kHz, according to the same procedure used with control group subjects. However, diseased group subjects did not have normal hearing. If a subject failed to respond to two presentations of a tonal stimulus at a given frequency, the tester moved on to testing other frequencies. Testers took measures to encourage subject cooperation to minimize the likelihood that an absent response was caused by a loss of interest in the task.

Middle ear condition was determined by a battery of test results including elevated air conduction thresholds, absent otoacoustic emissions, and abnormal immittance results using 226- and 1000-Hz probe tones.

Diseased Group Subjects Recruited through BCCH • Subjects tested through the BCCH Audiology Department had air and bone conduction thresholds measured at 0.5, 1, 2, and 4 kHz by a registered audiologist, and test results were required to have a good reliability rating. In addition to elevated air conduction thresholds, absent otoacoustic emissions, and abnormal impedance results, subjects recruited through the BCCH Otolaryngology Department had bone conduction thresholds established to confirm that the hearing loss was entirely conductive in nature. The diagnosis of middle ear pathology was confirmed based on pneumatic otoscopy and video otomicroscopy conducted by a pediatric otolaryngologist (OTL; N = 21 ears). Video otomicroscopy was later independently reviewed by an otologist to confirm the original diagnosis of MEE.

Statistical Analyses

A mixed-model analysis of variance (ANOVA) was used to analyze the control and diseased group data. The Greenhouse-Geisser (1959) approach was used in all cases to compensate for inflated type I error, which can exist among a large number of repeated measures. The alpha level was set at 0.05.

Experiment 1: Control Group Subjects

One-third octave averaging between 211 and 6000 Hz was used to reduce the number of frequencies from 248 to 15 for all statistical analyses. Normative pediatric ER data were measured on three different between-group factors: ethnicity (Caucasian versus Chinese), gender (male versus female), and ear (right versus left). The between-group factors were measured across 15 frequencies (repeated measures factor frequency with 15 levels). Normative pediatric ER data were then compared with the normative adult ER data obtained by Shahnaz and Bork (2006). ER was evaluated on factors of age (child versus adult) and ethnicity (Caucasian versus Chinese) across 15 frequencies.

Experiment 2: Diseased Group Subjects

Normative pediatric ER data were compared with ER from the three diseased groups across 15 frequencies on the between-group factor of middle ear condition (normal, mild negative middle ear pressure, severe negative middle ear pressure, and MEE).
will test the probability of the hypothesis that the difference in performance, as measured by AUROC values, is statistically significant. The AUROC curve can be used to compare the diagnostic performance of different measures or the same measure over different populations.

Post hoc analyses were performed to determine the frequency range over which significant differences existed between normative and MEE ER data. Receiver operating characteristic (ROC) curves were measured over this frequency range to statistically compare the diagnostic performance of ER as a function of frequency (Hilgers et al. 1990). ROC statistical analyses were performed using MedCalc for Windows, version 9.5.0.0 (MedCalc Software, Mariakerke, Belgium). The area under ROC (AUROC) curve and 95% confidence interval (CI) for this area were calculated automatically by the statistical software at each frequency value. The 95% CI is the interval in which the true (population) AUROC curve lies with 95% confidence (Hilgers 1991). Interpretation of the AUROC curve for a particular test frequency is as follows: the area value ranges between 0.5 and 1, where 0.5 indicates that ER at that frequency for distinguishing normal middle ear status from MEE is at a chance level and 1 indicates perfect test performance in distinguishing between the two middle ear conditions. An AUROC curve of 0.85 for 3400 Hz, for example, indicates that the ER value at 3400 Hz from an ear selected at random from the MEE group will be 85% of the time larger than the ER value at 3400 Hz from a normal middle ear ear (Zweig & Campbell 1993). The p value (as suggested by Hanley & McNeil 1982) was used to determine whether the AUROC is significantly different from 0.5, which provides evidence for the test’s ability to distinguish between the two groups.

The AUROC curve can also be used to compare the performance of two different measures or the same measure on two different occasions by comparing the statistical significance of the difference between the two AUROC curves (Hanley & McNeil 1983). This comparison does not rely on the selection of a particular decision threshold and will test the probability of the hypothesis that the difference between the two AUROC is zero. The method used will take into account the correlation between the areas that were obtained from the same sample of patients.

**RESULTS**

**Experiment 1: Control Group Subjects**

**Normative Pediatric Data** The main effects of ethnicity ($F_{1,116} = 0.31, p = 0.57$), gender ($F_{1,116} = 1.40, p = 0.23$), and age ($F_{1,116} = 2.04, p = 0.15$) were not significant, indicating that once ER was collapsed across the frequency range, it was not significantly different between Caucasian and Chinese, right and left, or between males and females. For these control group subjects, who ranged in age from 5 yrs 1 mo to 6 yrs 11 mo, there were no significant interactions between ethnicity, gender, and ear.

Following the G-G adjustment method, the interaction between frequency and ethnicity ($F_{1,1624} = 5.01, p = 0.00$) was the only significant interaction, indicating that ER varies as a function of frequency between the Caucasian and Chinese groups (Fig. 1). To investigate the frequencies at which the group differences occurred, a post hoc Tukey HSD test was performed. ER values in the Caucasian group were significantly higher (closer to 100%) than the Chinese group at 2000 and 6000 Hz.

**Pediatric Versus Adult ER Data** The normative pediatric ER data were compared with the normative adult ER data collected by Shahnaz and Bork (2006) (Figs. 2A, B). The data were explored using a mixed-model ANOVA. The main effects of age ($F_{1,357} = 1.61, p = 0.20$) and ethnicity ($F_{1,352} = 1.08, p = 0.29$) were not significant. The interaction between age and ethnicity was also not significant. Following the G-G adjustment model, the interactions between frequency and age ($F_{14, 4998} = 29.87, p = 0.000$), frequency and ethnicity ($F_{14, 4998} = 28.52, p = 0.000$), and frequency, age, and ethnicity ($F_{14, 4998} = 8.99, p = 0.000$) were significant. To investigate the frequencies at which the group differences occurred, a post hoc Tukey HSD test was performed. In the Caucasian group (Fig. 2A), children had significantly higher ER values (closer to 100%) than adults between 315 and 1250 Hz; however, ER values in adults were significantly higher than those measured in children between 2500 and 5000 Hz. In

---

**Fig. 1.** Mean energy reflectance as a function of frequency for Caucasian and Chinese children with normal middle ear status. Vertical bars denote 0.95 confidence interval.
the Chinese group (Fig. 2B), ER values in adults were significantly higher than children between 1250 and 2500 Hz. The outcome of the post hoc test was consistent with 0.95 CI plots of the mean in Figure 2, which shows no overlap between the two groups at these frequencies.

Experiment 2: Diseased Group Subjects
Data Analyzed by Middle Ear Condition • Normative pediatric ER data were compared with ER data from children with MEE, severe negative middle ear pressure, and mild negative middle ear pressure (Fig. 3). Given minimal differences between Caucasian and Chinese children, subject data from all ethnic groups, both genders and both ears were combined. ER consistently increases and increases over a wider range of frequencies as the middle ear condition changes from mild negative middle ear pressure to MEE. Also, the mean ER minimum occurred at approximately the same frequency (~3400 Hz) for all middle ear conditions.

Following G-G adjustment, the main effect of condition ($F_{3,236} = 180.53, p = 0.000$) was significant, indicating that ER varies differently between different middle ear conditions. The interaction between frequency and condition ($F_{42, 3304} = 22.84, p = 0.000$) was also significant, indicating that the variation of the ER across the frequency range was significantly different between different middle ear conditions. To investigate the frequencies at which the group differences occurred, a post hoc Tukey HSD test was performed. ER values measured from the group with MEE were significantly higher (closer to 100%) than ER values measured from the normal group between 680 and 6000 Hz. ER values measured from the group with significant negative middle ear pressure were significantly higher (closer to 100%) than ER values measured

---

**Fig. 2.** Mean energy reflectance as a function of frequency by age for Caucasian and Chinese subjects with normal middle ear status. Vertical bars denote 0.95 confidence interval.

**Fig. 3.** Mean energy reflectance (ER) as a function of frequency for different middle ear condition groups. Vertical bars denote 0.95 confidence interval.
from normal group between 630 and 5000 Hz. ER values measured from the group with mild and severe negative middle ear pressure were significantly higher (closer to 100%) than those measured from normal group from 630 to 2000 Hz and 3150 to 4000 Hz. ER values measured from the group with MEE were significantly higher than ER values measured from both groups with negative pressure between 1250 and 6000 Hz. ER values measured from the group with mild negative middle ear pressure were not significantly different from ER values measured from the group with severe negative middle ear pressure at any frequency tested.

**MEE Identified at Elementary Schools Versus OTL Confirmed at BCCH**

- Subjects identified through screenings at their elementary schools, who were suspected to have MEE based on elevated air conduction thresholds, flat tympanograms, and absent TEOAEs, were classified as “not OTL confirmed” MEE ($N = 21$). Subjects recruited through the BCCH Otolaryngology Department had MEE diagnosed by a pediatric OTL using pneumatic otoscopy and video otomicroscopy. These cases ($N = 21$) were classified as “OTL confirmed” MEE. A mixed-model ANOVA with frequency (15 levels) as a within-subject factor and method of documentation (OTL confirmed versus not OTL confirmed) as a between-subject factor was conducted. The main effect of MEE confirmation method was not significant ($F[1,40] = 0.04, p = 0.842$), and following the G-G adjustment, the interaction between frequency and MEE confirmation method was nonsignificant (Fig. 4). Data from both the OTL-confirmed and the not OTL-confirmed MEE group were combined for all analyses.

**Performance of ER in Distinguishing MEE from Normal Ears**

- Ethnic-specific analyses (Chinese only and Caucasian only) were performed in addition to analyses where data from all ethnic groups were combined. Although normative pediatric data reveal a significant interaction between ethnicity and frequency, using ethnic-specific normative data did not increase the test performance of ER in identifying MEE. Data from children of all ethnicities were combined in experiment 2 analyses.

Individual ER values for all 42 ears with MEE and for all 248 data points from 211 to 6000 Hz are shown in Figure 5 along with the normative mean ER and 80% range (10th to 90th percentile). This would result in a fixed false alarm (FA) rate of 20%, but because ER values always increase in the MEE compared with the control group, only the 90th percentile cutoff points were used for the differential diagnosis of MEE, setting the FA rate to 10%. The values of the MEE ears were all above the 90th percentile (ER closer to 100%) across a wide range of frequencies (~1200 to 2600 Hz), which means that the overall ER hit rate for distinguishing ears with MEE from normal ears was 100% and the FA rate was 10%.

Subsequent ANOVA using one-third octave band frequencies and post hoc analyses revealed significant differences between normative and MEE ER data from 630 to 6300 Hz. AUROC plots and corresponding 95% CI along with a sample of pair-wise comparisons of AUROC between different frequencies were automatically calculated by the statistical soft-
ware (MedCalc) and are summarized in Table 2. Only the data sample at 630 Hz is shown in Table 3, but identical analyses were conducted for all frequencies. Statistical comparison of the AUROC plots at multiple frequencies revealed that ER above 800 Hz (except for 6300 Hz) was able to distinguish normal middle ear status from MEE better compared with ER at 630 and 800 Hz. There was no statistical difference among AUROC curves measured from 1000 to 5000 Hz. ER at 1250 Hz was selected for further analysis because it had the largest AUROC (Table 2) and had better overall sensitivity and specificity. The ROC plot for ER at 1250 Hz is shown (Fig. 6), as the ER for 630 Hz and 6300 Hz to serve as a comparison. The dual plot showing the distribution of ER at 1250 Hz between normal ears and ear with MEE is shown in Figure 7. The horizontal line indicates the cutoff point with the best separation (maximum of the Youden index: sensitivity of 94.7% and specificity of 96.5%) between normal ears and ears with MEE. Similar procedures were also conducted for other middle ear conditions (negative middle ear pressures) and ER at 1250 Hz maintained the largest AUROC curve.

ER at 1250 Hz was statistically compared with static admittance ($Y_{tm}$) obtained automatically from the positive tail at conventional 226-Hz probe tone frequency using ROC plot analysis (Fig. 8). AUROC plots and corresponding 95% CI along with pair-wise comparison of AUROC between ER at

<table>
<thead>
<tr>
<th>Frequency</th>
<th>AUROC</th>
<th>SE</th>
<th>95% CI</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>ER_630</td>
<td>0.84</td>
<td>0.04</td>
<td>0.77–0.89</td>
<td></td>
</tr>
<tr>
<td>ER_800</td>
<td>0.93</td>
<td>0.03</td>
<td>0.88–0.96</td>
<td></td>
</tr>
<tr>
<td>ER_1000</td>
<td>0.96</td>
<td>0.02</td>
<td>0.93–0.99</td>
<td></td>
</tr>
<tr>
<td>ER_1250</td>
<td>0.97</td>
<td>0.02</td>
<td>0.94–0.99</td>
<td></td>
</tr>
<tr>
<td>ER_1600</td>
<td>0.97</td>
<td>0.02</td>
<td>0.93–0.99</td>
<td></td>
</tr>
<tr>
<td>ER_2000</td>
<td>0.95</td>
<td>0.02</td>
<td>0.91–0.98</td>
<td></td>
</tr>
<tr>
<td>ER_2500</td>
<td>0.95</td>
<td>0.02</td>
<td>0.90–0.97</td>
<td></td>
</tr>
<tr>
<td>ER_3150</td>
<td>0.95</td>
<td>0.02</td>
<td>0.91–0.98</td>
<td></td>
</tr>
<tr>
<td>ER_4000</td>
<td>0.96</td>
<td>0.02</td>
<td>0.91–0.98</td>
<td></td>
</tr>
<tr>
<td>ER_5000</td>
<td>0.92</td>
<td>0.03</td>
<td>0.87–0.95</td>
<td></td>
</tr>
<tr>
<td>ER_6300</td>
<td>0.89</td>
<td>0.04</td>
<td>0.84–0.93</td>
<td></td>
</tr>
</tbody>
</table>

AUROC, area under the receiver operating characteristic curve; CI, confidence interval; ER, energy reflectance; SE, standard error.

1250 Hz and $Y_{tm}$ at 226 Hz are summarized in Tables 4 and 5. Statistical comparison of the AUROC plots among these measures revealed that with respect to distinguishing ears with MEE from normal ears, ER at 1250 Hz was superior to conventional 226 Hz measure of $Y_{tm}$.

**Test-Retest Reliability** • Analyses were performed to ensure that ER differences observed between the MEE subjects and the control subjects were caused by a difference in middle ear condition rather than variability in test measurement techniques, such as reinsertion of the ear tips. Thirty-five ears from the control group were randomly selected to serve as an estimate of probe tip reinsertion test-retest reliability. After the collection of the WBR data, the probe was removed for 1 to 2 mins and subsequently reinserted into the same ear canal. ER was remeasured after the reinsertion. The difference between the initial and secondary insertion ER values was calculated at each of the frequencies, similar to the mean insertion difference.

To serve as a comparison, ER measures of 10 healthy ears and 10 MEE ears were selected at random. For each of the 10 cases, the ER values of the healthy ears were subtracted from the ER values of those with effusion at each of the frequencies. A mean value was calculated (Fig. 9). The mean insertion

<table>
<thead>
<tr>
<th>Frequency</th>
<th>AUROC</th>
<th>SE</th>
<th>95% CI</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>ER_630 vs. ER_800</td>
<td>0.09</td>
<td>0.02</td>
<td>0.05–0.13</td>
<td>0.00</td>
</tr>
<tr>
<td>ER_630 vs. ER_1000</td>
<td>0.13</td>
<td>0.03</td>
<td>0.07–0.18</td>
<td>0.00</td>
</tr>
<tr>
<td>ER_630 vs. ER_1250</td>
<td>0.14</td>
<td>0.04</td>
<td>0.06–0.21</td>
<td>0.00</td>
</tr>
<tr>
<td>ER_630 vs. ER_1600</td>
<td>0.13</td>
<td>0.04</td>
<td>0.05–0.21</td>
<td>0.00</td>
</tr>
<tr>
<td>ER_630 vs. ER_2000</td>
<td>0.11</td>
<td>0.04</td>
<td>0.03–0.19</td>
<td>0.01</td>
</tr>
<tr>
<td>ER_630 vs. ER_2500</td>
<td>0.11</td>
<td>0.05</td>
<td>0.02–0.19</td>
<td>0.01</td>
</tr>
<tr>
<td>ER_630 vs. ER_3150</td>
<td>0.11</td>
<td>0.05</td>
<td>0.02–0.21</td>
<td>0.01</td>
</tr>
<tr>
<td>ER_630 vs. ER_4000</td>
<td>0.12</td>
<td>0.05</td>
<td>0.03–0.21</td>
<td>0.01</td>
</tr>
<tr>
<td>ER_630 vs. ER_5000</td>
<td>0.08</td>
<td>0.05</td>
<td>−0.01 to 0.17</td>
<td>0.08</td>
</tr>
<tr>
<td>ER_630 vs. ER_6300</td>
<td>0.05</td>
<td>0.05</td>
<td>−0.04 to 0.15</td>
<td>0.27</td>
</tr>
</tbody>
</table>

AUROC, area under the receiver operating characteristic curve; ER, energy reflectance; SE, standard error; CI, confidence interval.

Fig. 7. A dual dot plot for energy reflectance (ER) at 1250 Hz with a horizontal line indicating the cutoff point with the best separation (minimal false-negative and false-positive results) between normal ears and ears with otitis media with effusion. Sens, sensitivity; Spec, specificity.
difference is negligible compared with the difference that was measured between children with normal ears and those with MEE.

**Ear-specific Follow-Up** • There was a small group of ears (N = 4) that were tested through elementary schools and were classified as MEE (not OTL confirmed) where a follow-up screening occurred 6 wks later. In all four cases, audiological test battery results were normal at the time of the reassessment, and MEE was no longer suspected. An example is illustrated in Figure 10, where initial and follow-up ER data are represented along with the 90% range of the normative pediatric data. Note that a much larger proportion of the ER pattern falls within the 90% range at the time of the follow-up test compared with ER measured on the initial test date.

**DISCUSSION**

The goals of this study were to establish normative ER data for early school-aged children; to determine whether ER differs significantly between Caucasian and Chinese children, male and female children, and children and adults (experiment 1); and to compare the normative pediatric ER data with the ER data obtained from children with abnormal middle ear conditions (experiment 2). To our knowledge, there is no published ER data of this sort.

**Experiment 1: Control Group Data**

**Effects of Ear, Gender, Ethnicity, and Age** • The effects of ear and gender and their interactions with frequency were not significant. This is consistent with findings of Hunter et al. (2008) in newborns and toddlers and justifies pooling of ER data from right and left, and male and female ears in this study. However, Keefe et al. (2000) found both ear and gender differences in a neonatal population, which is not consistent with our findings. They reported that “the left ear was acoustically stiffer than the right ear, and that the female ear was acoustically stiffer than the male ear.” The source of these ear and gender differences is not known but may be attributable to subject age, subject age range, or experiment methodology. It should be noted that Keefe et al. reported ear and gender differences from a larger sample size, and it is possible that this study did not have sufficient participant numbers to detect this small difference.

ER varied as a function of frequency between Caucasian and Chinese children and between children and adults. The ER differences that existed within an adult population between male and female and Caucasian and Chinese ears (Shahnaz & Bork 2006) were much more pronounced compared with ER differences measured in this study within a pediatric population. Shahnaz and Bork explained that adult ER differences could be partly attributed to differences in body size between the groups (Bell et al. 2002). Although data comparing height, weight, and body mass index among children of different ethnic groups were not available, minimal differences in body size indices between Caucasian and Chinese children may explain the smaller differences observed between Caucasian and Chinese ER patterns in this study compared with those observed within an adult population (Shahnaz & Bork 2006).

Significantly higher ER values (closer to 100%) were measured in Caucasian children compared with Caucasian adults between 315 and 1250 Hz, and significantly higher ER values were measured in both Caucasian and Chinese adults.
compared with Caucasian and Chinese children between 2500 and 5000 Hz. This suggests that the adult middle ear is capable of more sound absorption at low frequencies, whereas the pediatric middle ear is capable of more sound absorption at higher frequencies. An increase in body size in animal models is associated with an increase in the size of their middle ear structures, such as increased ossicle and footplate size and increased tympanic membrane area (Werner et al. 1998; Werner & Igic 2002), which increases the mass of the conductive mechanism and can degrade the high-frequency response of the middle ear (Relkin 1988; Saunders et al. 1998). It is unclear how ER relates to the mass of the conductive mechanism; however, it could partially explain higher ER values at high frequencies in the adult group compared with the pediatric group. This is also consistent with multifrequency tympanometric data that have shown higher middle ear resonant frequency in children aged 6 to 15 yrs (Hanks & Rose 1993) compared with adults (Shanks et al. 1993; Shahnaz & Polka 1997). However, it should be noted that WBR measurements alone cannot provide a direct assessment of the forward energy transfer through the middle ear and into the cochlea.

Some of the observed ER differences between adults and children may be explained by a middle ear model proposed by Feeney and Sanford (2004) which consists of a serial combination of resistance, stiffness, and mass. With stiffness decreased by a known factor in the elderly group, the model was able to explain ER differences observed between a group of young and elderly adults. The elderly group had lower ER than the young group at frequencies below the reflectance minimum (the frequency at which ER is closest to zero) and higher ER at frequencies above the reflectance minimum. The adult group in this study, similar to the elderly group in the study by Feeney and Sanford, had lower ER than the pediatric group at frequencies below the reflectance minimum and higher ER at frequencies above the reflectance minimum.

Experiment 2: Diseased Group Subjects

ER Data for Different Middle Ear Conditions • Data from this experiment reveal that between normal middle ear status and a mild degree of negative pressure, changes in energy absorption are most evident over the low-frequency range (from ~400 to 1800 Hz). This may be attributable to the increased stiffness of the middle ear system and is consistent with the findings of Margolis et al. (2001) who demonstrated changes in ER patterns caused by induced pressure changes within the ear canal. More recently, Hunter et al. (2008) showed that increased or decreased tympanometric peak pressure results in an ER increase, primarily below 1000 Hz.

Between middle ear states of negative pressure and effusion, the presence of fluid within the middle ear cavity increases the mass and the stiffness of the middle ear system. ER over the low-frequency range is maximal with negative middle ear pressure, and the most notable increase in ER between the negative pressure and effusion conditions is over the mid- to high-frequency range (from ~1000 to 6000 Hz). The measured increase in ER at higher frequencies may be a direct result of the increased mass load on the middle ear system. This is consistent with the findings of Hunter et al. (2008) who demonstrated a significant difference in middle ear reflectance from 1 to 4 kHz in children 3 days to 47 mo of age with clinically defined OME. Moreover, Voss et al. (2008) found that variations in middle ear cavity volume can largely affect ER, especially below 1000 Hz (larger ER values, closer to 1, exist in smaller cavities). This is consistent with current findings because the presence of fluid reduces middle ear cavity volume.

Test Performance and Test-Retest Reliability of ER • In all 42 ears with MEE, ER was greater than the 90th percentile range of the normative data, particularly for frequencies greater than 600 Hz. The pattern of how ER changed from normal was consistent across all ears with MEE. This increase in ER is consistent with previously reported ER data in smaller samples of ears with MEE (Feeney et al. 2003; Allen et al. 2005).

The overall test performance of ER was objectively evaluated using ROC analyses and was compared across frequencies averaged in one-third octave bands (15 frequencies). Overall, test performance of ER averaged in one-third octave bands was comparable between 1000 and 5000 Hz. ER at 1250 Hz had the...
largest AUROC (sensitivity of 96% and specificity of 95%) and was selected for further analysis. Comparison of AUROC between ER at 1250 Hz and tympanometric static admittance obtained at 226 Hz revealed significantly better test performance for ER in distinguishing between MEE and normal middle ear status.

ER measures were found to reflect true differences in middle ear condition rather than a source of test error, such as probe tip reinsertion. Small test-retest variability suggests that test-retest differences were not the source of differences in middle ear conditions.

Limitations of the Study (Experiments 1 and 2)

It is not known whether the effect of age confounded the WBR data of the diseased group subjects. The age range of the normal subjects and the diseased group subjects who were recruited through the elementary schools was strictly limited to 5- and 6-yr-old children, but patients recruited through the BCCH Otolaryngology Department were between the ages of 3 and 12 yrs. A looser age criterion was permitted to increase the number of pathological ears tested.

The Mimosa calibration method is based on that of Allen (1986) and Voss and Allen (1994) (Jeng, Reference Note 1). Impedance measured by the Mimosa instrument is divided by characteristic impedance to obtain normalized values and attempt to compensate for the effect of the ear canal using constant values (Voss & Allen 1994). If these constants, such as cross-sectional ear-canal area, are different between school-aged children and adults, then any observed effect of age may partly include an effect of cross-sectional ear-canal area on the ER calculation. Larger ear-canal diameters could potentially result in lower ER values (Sanford & Feeney 2008) as observed between Caucasian adults and children. Although there is some evidence that cross-sectional area of the ear canal in newborns changes as a function of age because of maturation (Keefe & Abdala 2007), there are no published studies that have shown differences in cross-sectional ear-canal area between school-aged children and adults. This is a minor issue in a clinically directed investigation, where middle ear dysfunction is assumed to make larger shifts in WBR, but may be more noteworthy when looking to detect smaller shifts in WBR, such as WBR differences observed between ethnic groups.

Implications of the Study

WBR shows promise as a clinical diagnostic tool for middle ear analysis. It is able to generate frequency-specific measures of middle ear acoustic transfer and is sensitive to changes in middle ear status. The mechaanoacoustical properties of a healthy middle ear system differ between pediatric and adult populations and also between Caucasian and Chinese children. However, the ER differences observed between Caucasian and Chinese children were not large enough to affect the test performance of ER in identifying middle ear pathology.

Directions for Future Research

WBR measures must be evaluated in terms of sensitivity and specificity against the other available measures of middle ear status, such as multifrequency tympanometry, to determine which middle ear measure or combination of measures are able to provide information with the highest test performance. This could determine which middle ear assessment methods should be used clinically and may alter the methods used for screening school-aged children.

Once reliable normative data are established, further experimentation can be conducted to determine sources of WBR differences among patients with MEE. WBR may have the ability to identify the volume and/or viscosity of MEE. WBR may become a useful clinical and preoperative tool for OTRs if it is able to successfully predict the nature of middle ear fluid noninvasively.

CONCLUSIONS

A preliminary set of normative ER data have been generated for a pediatric population between the ages of 5 and 7 yrs, and it differs significantly from normative adult ER data gathered previously by Shahnaz and Bork (2006). In this study, pediatric normative data were warranted for testing children, but race-specific norms were not required to detect middle ear pathology and changes in middle ear status. WBR shows promise as a clinical diagnostic tool for measuring the mechanoaoustic properties of the middle ear and the changes that result in the presence of negative middle ear pressure and/or MEE.

ACKNOWLEDGMENTS

The authors thank Dr. Shahnaz Atashband and Ms. Michele Freeman for their assistance with recruiting and scheduling subjects, Ms. Laurie Usher and the BC Children’s Hospital Audiology Department for their support of the project, the Richmond and Coquitlam school districts for their cooperation, and all of the students (and their families) who participated in the study. The authors also thank Ms. Nerissa Davies and Ms. Vahideh Boshaghzadeh for their assistance with subject testing.

This project was financially supported by the Canadian Foundation for Innovation (CFI) and the British Columbia Knowledge Development Fund (BCKDF).

Preliminary data from this study were presented at the Annual Convention of the American Auditory Society and the Annual Meeting of the Society for Ear, Nose, and Throat Advances in Children in 2008.

Part of this research is based on a Masters thesis by the first author under the supervision of the second author.

Address for correspondence: Navid Shahnaz, Faculty of Medicine, School of Audiology and Speech Sciences, University of British Columbia, 2177 Wesbrook Mall, Vancouver, British Columbia, Canada V6T 1Z3. E-mail: nshahnaz@audiospeech.ubc.ca.

Received November 10, 2008; accepted August 25, 2009.

REFERENCES


REFERENCE NOTE