J Am Acad Audiol 29:106–117 (2018)

Predictive Accuracy of Sweep Frequency Impedance Technology in Identifying Conductive Conditions in Newborns

DOI: 10.3766/jaaa.16077

Venkatesh Aithal*†
Joseph Kei†
Carlie Driscoll†
Michio Murakoshi‡
Hiroshi Wada§

Abstract

Background: Diagnosing conductive conditions in newborns is challenging for both audiologists and otolaryngologists. Although high-frequency tympanometry (HFT), acoustic stapedial reflex tests, and wideband absorbance measures are useful diagnostic tools, there is performance measure variability in their detection of middle ear conditions. Additional diagnostic sensitivity and specificity measures gained through new technology such as sweep frequency impedance (SFI) measures may assist in the diagnosis of middle ear dysfunction in newborns.

Purpose: The purpose of this study was to determine the test performance of SFI to predict the status of the outer and middle ear in newborns against commonly used reference standards.

Research Design: Automated auditory brainstem response (AABR), HFT (1000 Hz), transient evoked otoacoustic emission (TEOAE), distortion product otoacoustic emission (DPOAE), and SFI tests were administered to the study sample.

Study Sample: A total of 188 neonates (98 males and 90 females) with a mean gestational age of 39.4 weeks were included in the sample. Mean age at the time of testing was 44.4 hr.

Data Collection and Analysis: Diagnostic accuracy of SFI was assessed in terms of its ability to identify conductive conditions in neonates when compared with nine different reference standards (including four single tests [AABR, HFT, TEOAE, and DPOAE] and five test batteries [HFT + DPOAE, HFT + TEOAE, DPOAE + TEOAE, DPOAE + AABR, and TEOAE + AABR]), using receiver operating characteristic (ROC) analysis and traditional test performance measures such as sensitivity and specificity.

Results: The test performance of SFI against the test battery reference standard of HFT + DPOAE and single reference standard of HFT was high with an area under the ROC curve (AROC) of 0.87 and 0.82, respectively. Although the HFT + DPOAE test battery reference standard performed better than the HFT reference standard in predicting middle ear conductive conditions in neonates, the difference in AROC was not significant. Further analysis revealed that the highest sensitivity and specificity for SFI (86% and 88%, respectively) was obtained when compared with the reference standard of HFT + DPOAE. Among the four single reference standards, SFI had the highest sensitivity and specificity (76% and 88%, respectively) when compared against the HFT reference standard.

Conclusions: The high test performance of SFI against the HFT and HFT + DPOAE reference standards indicates that the SFI measure has appropriate diagnostic accuracy in detection of conductive conditions in newborns. Hence, the SFI test could be used as adjunct tool to identify conductive conditions in universal newborn hearing screening programs, and can also be used in diagnostic follow-up assessments.

*Audiology Department, Townsville Hospital and Health Service, Queensland, Australia; †Hearing Research Unit for Children, School of Health and Rehabilitation Sciences, University of Queensland, Brisbane, Australia; ‡Biomechanical Engineering Laboratory, Department of Mechanical Engineering, Kagoshima University, Kagoshima, Japan; §Department of Intelligent Information System, Tohoku Bunka Gakuen University, Sendai, Japan

Corresponding author: Venkatesh Aithal, Audiology Department, Townsville Hospital and Health Service, Queensland, 4814, Australia; E-mail: venkatesh.aithal@health.qld.gov.au

This research was funded by the Healthy Hearing Program, Private Practice Research and Education Trust fund, QH Health Practitioners Research fund, and thesis writing grant.

Key Words: dynamic behavior, high-frequency tympanometry, middle ear, multifrequency tympanometry, newborn, reference standards, resonance frequency, ROCs, sweep frequency impedance, test performance

Abbreviations: AABR = automated auditory brainstem response; AROC = area under receiver operating curve; daPa = deca Pascal; DPOAE = distortion product otoacoustic emissions; HFT = high-frequency tympanometry; LR = likelihood ratio; OAE = otoacoustic emissions; OME = otitis media with effusion; Ps = static pressure; RF = resonance frequency; ROC = receiver operating characteristics; SD = standard deviation; SFI = sweep frequency impedance; SN = sensorineural; SNR = signal-to-noise ratio; SPL = sound pressure level; TEOAE = transient evoked otoacoustic emissions; UNHS = universal newborn hearing screening; WBA = wideband absorbance; WBR = wideband reflectance; Ya = static admittance; Ypc = peak compensated static admittance

INTRODUCTION

Universal newborn hearing screening (UNHS) to detect permanent hearing loss using otoacoustic emissions (OAE) or automated auditory brainstem response (AABR) has become standard practice across many countries. Despite improvements in the screening technology, high rates of referrals due to transient conductive conditions continue to be an issue with UNHS programs. A child who failed the screening test but is later identified to have normal hearing is regarded as having a transient conductive hearing loss (Clemens et al, 2000; Clemens and Davis, 2001; Mehl and Thomson, 2002; Keefe and Feeney, 2009). However, some children may be found to have a congenital or long-standing conductive condition which requires medical intervention (Boudewyns et al, 2011). Several studies have reported high rates of referrals of 3–8% in newborns (Mason and Hermann, 1998; Mehl and Thomson, 1998; Vohr et al, 1998; Clemens et al, 2000; Clemens and Davis, 2001) that is attributed to transient conductive loss due to outer and middle ear dysfunction. As expected, middle ear dysfunction or cochlear hearing loss obliterates OAE and AABR responses, resulting in a "refer" outcome. It is, therefore, not possible to distinguish between a conductive or cochlear hearing loss when a refer outcome occurs (Allen et al, 2005).

In newborns, a transient conductive condition may be due to vernix occluding the ear canal, residual amniotic fluid or mesenchyme in the middle ear space of well neonates (Buch and Jorgensen, 1964; Kok et al, 1992; Keefe et al, 2000; Rosenfeld et al, 2004) and neonates cared for in the neonatal intensive care unit (Balkany et al, 1978; Paradise, 1981; Derkay et al, 1988). Transient outer and middle ear dysfunction is common in neonates tested within 48 hr of birth and infants in the neonatal intensive care unit for extended periods of time (Keefe and Feeney, 2009). As most newborns are screened within first 48 hr of birth, temporary conductive conditions may contribute to high referral rates (Doyle et al, 2000; Allen et al, 2005).

Many studies have reported the prevalence of conductive hearing loss due to transient middle ear dysfunction such as otitis media with effusion (OME) to be higher than sensorineural (SN) hearing loss (Doyle et al, 2000; Keefe et al, 2000; Boone et al, 2005; Hunter et al, 2010; Silverman, 2010; Aithal et al, 2012). For example, in a retrospective study of 76 infants, Boone et al (2005) attributed 64.5% of failures in UNHS to OME. In another follow-up study of 211 infants referred for diagnostic assessment following AABR screening, 32% had a conductive loss due to middle ear dysfunction (Aithal et al, 2012). Similarly, Boudewyns et al (2011) reported 55.3%, Holster et al (2009) reported 18.4%, and Doyle et al (2004) reported 58% of failures due to OME. These studies also noted that OME is an important cause of transient hearing loss during the first months of life (Doyle et al, 2004; Holster et al, 2009; Boudewyns et al, 2011). These studies also highlighted that infants who failed in UNHS due to OME are at increased risk for later development of chronic otitis media.

Despite the high referral rates due to middle ear dysfunction, there is presently no single validated objective tool that can be used at the time of screening to assess middle ear function (Keefe et al, 2000). Assessment of middle ear function is not currently part of the JCIH (2007) guidelines for UNHS programs. Middle ear assessment is only recommended for newborns as part of the diagnostic assessment. Hence, differential diagnosis of transient conductive loss and permanent SN loss can only be made during follow-up diagnostic assessments, which are expensive and time consuming. For this reason, development of screening tools to assess middle ear function at the time of newborn screening is recommended (Gravel et al, 2005). Such tools would assist in streamlining the management strategies for the respective types of hearing loss, facilitating prioritization of neonates for follow-up appointments, and reducing parental anxiety.

The only definitive tests for the presence of OME are myringotomy or imaging studies such as computerized axial tomography scanning. However, neither of these techniques is practical or ethical under screening conditions. Otoscopy is not reliable in newborns due to difficulties in observing changes in color, mobility, reflexive reaction to lights, and translucency of the ear drum (Paradise, 1980; Ruah et al, 1991; Rhodes et al, 1999) and it cannot be relied upon in the diagnosis of OME in newborns (Shurin

et al, 1976; Roberts et al, 1995). For example, Shurin et al (1976) noted that five of ten ears diagnosed with OME by otoscopy were found to have normal or dry ear on tympanocentesis. Similarly, Doyle et al (1997) observed that half of the 9% of ears diagnosed as having OME based on reduced ear drum mobility on pneumatic otoscopy passed the AABR screening test and about one-third passed the transient evoked OAE (TEOAE) screening test, indicating that the use of pneumatic otoscopy in young infants can result in incorrect diagnosis.

Previous attempts using conventional 226-Hz tympanometry to diagnose middle ear dysfunction in young infants (≤6 mo of age) have been unsuccessful (Paradise, 1976; Keefe and Levi, 1996; Rhodes et al, 1999; Purdy and Williams, 2002; Baldwin, 2006).The use of high-frequency tympanometry (HFT) with a 1000-Hz probe tone has been reported to be more successful (Purdy and Williams, 2002; Kei et al, 2003; Margolis et al, 2003; Baldwin, 2006; Swanepoel et al, 2007). The HFT is currently recommended by JCIH (2007) for diagnostic testing after UNHS referrals. The sensitivity and specificity of HFT in detecting conductive conditions in newborns using distortion product OAE (DPOAE) as a reference standard are reported to be 50% and 91%, respectively (Margolis et al, 2003). While Swanepoel et al (2007) reported a similar test performance result of sensitivity of 57% and specificity of 95% for HFT against a DPOAE reference standard, Baldwin (2006) reported a sensitivity of 99% and specificity of 89% using AABR outcomes as the reference standard for older infants with a mean age of 10 weeks. However, Baldwin's findings might not apply for neonates as the youngest infant in her study was two weeks old. Although HFT is recommended for use with young infants (Kei et al, 2003; Margolis et al, 2003; Baldwin, 2006; Alaerts et al, 2007), there are no universally agreed methods for interpreting results (Kei and Mazlan, 2012). It is also shown that introducing negative and positive air pressure distends an infant ear canal and modifies the middle ear characteristics (Holte et al, 1990). It also violates the underlying assumptions of tympanometry in infants (Margolis and Shanks, 1990; Kei and Zhao, 2012). For instance, on pressurization as in tympanometry, the diameter of ear canal increases by an average of 18.3% under positive pressure or decreases by 28.2% under negative pressure compared to its original value (Holte et al, 1990). Furthermore, ear canal volume changes from 27% to 75% over a range of ±300 deca Pascal (daPa) in newborns (Qi et al, 2006). In addition to these limitations, measurement of peak-compensated static admittance (Ypc) for the negative tail method introduces artifactual spikes and danger of collapsing the ear canal with negative ear canal pressure (Kei et al, 2007; Aithal et al, 2016; Hunter and Blankenship, 2017). It is also reported that the positive tail method overestimates ear canal volume and is less sensitive to middle ear dysfunction in young infants (Hunter and Blankenship, 2017). Given such limitations of HFT, there is a need to introduce alternative techniques to identify middle ear dysfunction in young infants.

Wideband absorbance (WBA) is an emerging technology that can assess middle ear function in young infants and could be a useful tool in UNHS (Keefe et al, 2003; Sanford et al, 2009; Hunter et al, 2010; Hunter et al, 2015). However, it is not yet recommended by JCIH (2007) for diagnostic testing of infants. Recently, Aithal et al (2015a) attempted to use WBA technology to identify conductive conditions in newborns. The authors evaluated the test performance of WBA against commonly used reference standards using receiver operating characteristics (ROC) curve analysis. Their results showed that optimal test performance of the WBA, as indicated by the area under the ROC curve (AROC), reached 0.78 when compared against test battery reference standards (Aithal et al, 2015a). Sanford et al (2009) and Hunter et al (2010) reported better performance of WBA than HFT in predicting conductive conditions in newborn screening whereas other studies have reported that both HFT and WBA are good measures to identify transient conductive hearing loss in young infants (Sanford et al, 2012; Prieve et al, 2013b). Sanford et al (2009) reported the highest AROC of 0.86 (95% confidence interval [CI]: 0.80–0.89) for WBA and 0.75 (95% CI: 0.68–0.80) for HFT with DPOAE screening as the reference standard on day one. On day two, they reported an AROC of 0.67 (95% CI: 0.45–0.83) for WBA and 0.54 (95% CI: 0.36–0.71) for HFT. Similarly, Hunter et al (2010) reported that wideband reflectance (WBR) produced much better prediction of DPOAE status in newborns than HFT and reported an AROC of 0.72 for HFT, 0.82 for WBR at 1 kHz, and 0.90 for WBR at 2 kHz. Although WBA performed better than HFT, both measures were proven to be effective in detecting conductive conditions in neonates (Sangster, 2011).

An alternative measure of outer and middle ear function, the sweep frequency impedance (SFI) technology, has been found to be more accurate than tympanometry in diagnosing middle ear diseases in adults (Wada and Kobayashi, 1990; Wada et al, 1998). However, the application of SFI to detect conductive conditions in young infants has not been investigated until recently (Murakoshi et al, 2012; 2013; Aithal et al, 2014; 2015b; 2016). The SFI method, developed in the 1990s, measures the resonance frequency (RF) and mobility (ΔSPL) of the outer and middle ear system at different static pressures in the ear canal as well as tympanometric peak pressure (the pressure at which the sound pressure level (SPL) attains maximum value) (Wada and Kobayashi, 1990; Murakoshi et al, 2012; Zhao and Wang, 2012). Although the technology appears to measure impedance, it actually measures the SPL in the ear canal while a sweeping tone is presented under various static pressure levels in the ear canal. From these SPL curves, the dynamic behavior of the

outer and middle ear can be described in a graph showing the SPL (in dB SPL) against frequencies from 100 to 2200 Hz at various static pressures applied to the ear canal. From the SPL curves, the RF and ΔSPL can be measured.

The dynamic behavior of the outer and middle ear system, as measured using an SFI meter, of a healthy newborn is shown in Figure 1A. The greatest variations of SPL (ΔSPL) occur at median frequencies of RF1 and RF2 along the SPL curve (at ambient pressure of 0 daPa) which are considered the RF of the ear canal and middle ear, respectively (Wada et al, 1995; Murakoshi et al, 2013). In comparison, the SFI results for a newborn, who failed in HFT and TEOAE, reveal only one variation at RF1 (Aithal et al, 2016) (Figure 1B).

Previous studies have demonstrated that the SPL curves obtained from neonates have shown two variations, that is, low-frequency (210–420 Hz) and high-frequency (830–1500 Hz) regions (Murakoshi et al, 2013; Aithal et al, 2014). This suggests that there are two vibrating elements in the neonatal auditory system, possibly due to external and middle ear components. The second variation in the higher frequency region (RF2) was similar to that of adult middle ear RF reported in the Wada et al (1998) study.

The first variation in the low-frequency region (RF1) was thought to be caused by an element other than the middle ear. It was considered to be associated with the movement of the external ear canal wall as Young's modulus of which is estimated to be 0.36 times as much as that of adults, that is, 36–364 kPa (Saunders et al, 1983; Qi et al, 2006). The resonance movements of the neonatal ear canal have been reported to be <450 Hz (Keefe et al, 1993). In addition, studies based on a neonatal external ear canal physical model using agarose gel and a numerical model using the finite element method suggest that the external ear canal wall exhibits intrinsic oscillatory behavior at ~300 Hz (Murakoshi et al, 2013; Hamanishi et al, 2015). Hence, the first variation of the SPL curve obtained from neonates in the low-frequency region (RF1) is considered to be related to the resonance of the neonatal external ear canal wall movements (Murakoshi et al, 2013; Hamanishi et al, 2015; Wada et al, 2016).

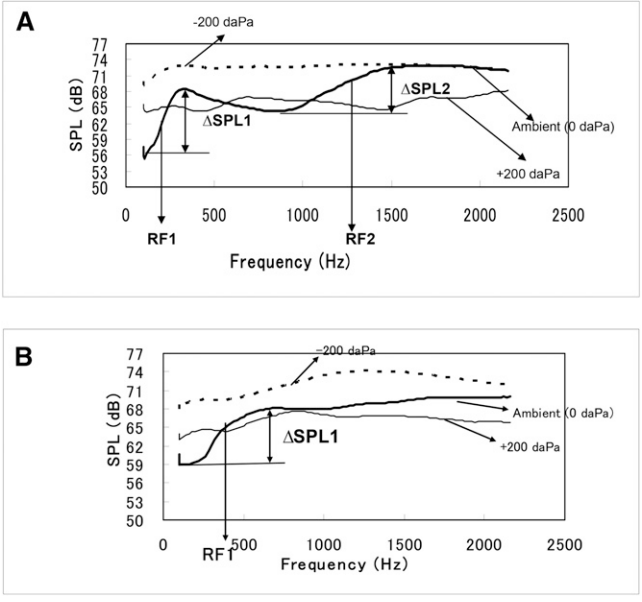

Figure 1. Typical SFI results obtained (A) from a healthy one-day-old newborn who passed the test battery. The SPL curve at ambient pressure shows two inflections in sound pressure (RF1 and RF2). The greatest variations of SPL (ΔSPL) occur at the median frequencies of RF1 and RF2 which are considered the RF of the ear canal and middle ear, respectively, (B) from a one-day-old newborn who did not pass the test battery. Note: The SFI curve at ambient pressure shows only one inflection (RF1). The static ear canal pressures (daPa) applied were +200, 0 (ambient), and −200 daPa.

As illustrated in Figure 1B, Murakoshi et al (2012) and Aithal et al (2016) noted that the second resonance (RF2) disappeared in newborns with middle ear dysfunction while the first resonance (RF1) was not affected. These results indicated that middle ear dysfunction altered the dynamic behavior of the middle ear to such an extent that the second resonance could not be detected using the SFI meter. SFI measures have potential advantages over HFT in assessing infants. First, since pressurization of the ear canal is not required when SFI measures are obtained at ambient pressure, distortion or collapse of the ear canal wall is not a concern. Second, measurements are made over a wide frequency range from 100 to 2200 Hz, rather than at a single frequency, and finally, SFI can provide additional information such as RF and mobility of the middle ear which, may assist in diagnosing conductive conditions in young infants (Murakoshi et al, 2012; Aithal et al, 2016). While SFI shows promising results in identifying dysfunction of the outer and middle ear in newborns, the test performance of SFI in determining the middle ear status of newborns has not been evaluated against any reference standard. The predictive accuracy of SFI must be compared to other reference standards before the SFI can be used as a mass screening tool for identifying conductive conditions in newborns. The aim of this study was, therefore, to evaluate the test performance of SFI to predict the middle ear status in newborns against clinical reference standards including four single tests (AABR, HFT, DPOAE, and TEOAE) and their combinations (HFT + DPOAE, HFT + TEOAE, DPOAE + TEOAE, DPOAE + AABR, and TEOAE + AABR).

MATERIALS AND METHODS

This study was approved by the Human Research Ethical Committee of Townsville Hospital and Health Service, and the University of Queensland Behavioral and Social Sciences Ethical Review Committee.

Parents provided written consent for newborns to be included in the study. All infants were born at full term, with normal birth weight and no medical complications or risk factors for hearing loss.

In total, 188 (98 males and 90 females) healthy neonates were recruited for the present study. All measurements were performed in a quiet room in the maternity unit where the noise level was <40 dBA. Nursing staff performed the AABR screen on both ears of all neonates as part of a state-mandated UNHS. AABR screening was performed using an ALGO3 newborn hearing screener (Natus Medical Inc., Pleasanton, CA). Clicks were presented at 35 dB nHL to both ears simultaneously during testing. A pass or refer result for each ear was automatically recorded by the equipment.

An experienced audiologist conducted HFT, DPOAE, TEOAE, and SFI tests on the neonates with the accessible ear tested first, followed by the second ear if possible. HFT was performed using a Madsen Otoflex 100 acoustic immittance device (GN Otometrics, Taastrup, Denmark) with a 1000-Hz probe tone of 75 dB SPL delivered to the ear (Mazlan et al, 2009). Admittance (Ya) was measured as the pressure was changed from +200 to −400 daPa at a rate of 400 daPa/sec. Pass criteria were a single positively peaked tympanogram with middle ear pressure between 50 and −150 daPa and peak-compensated static admittance Ypc (+200 daPa tail to peak) of ≥0.2 mmho (Mazlan et al, 2009)

DPOAE testing was performed using a Scout Sport (Biologic Navigator Plus, Mundelein, IL) system. DPOAEs were obtained in response to stimulation by pairs of primary tones. The f2/f1 frequency ratio was 1.2 for each primary pair. The level of f1 was 65 dB SPL and f2 was 55 dB SPL. The pass criteria included the following: (a) DPOAE-to-noise ratio of at least 6 dB in at least three of four frequencies from 2 to 6 kHz (Sanford et al, 2009; Hunter et al, 2010) and (b) DPOAE amplitude of at least −6 dB SPL at 2, 3, 4, and 6 kHz (Sanford et al, 2009; Merchant et al, 2010).

TEOAE testing was conducted using the same device mentioned above. Wideband clicks of 80-μsec duration were delivered to the ear at 80 dB peak SPL. Emissions were measured at 1, 1.5, 2, 3, and 4 kHz. The pass criteria included reproducibility of ≥70% and a signal-to-noise ratio (SNR) of ≥3 dB at 2, 3, and 4 kHz (Kei et al, 2003; Vander Werff et al, 2007).

The SFI test was performed using a new SFI unit developed for testing neonates (Murakoshi et al, 2013; Aithal et al, 2014). The SFI unit and its calibration have been described in detail by Murakoshi et al (2013) and Aithal et al (2014). A brief description of the SFI unit is provided here. Figure 2 shows a block diagram of the SFI unit that consisted of a personal computer, an analog-to-digital/digital-to-analog converter, a probe system, a stepping motor, an air pump, a pressure sensor, and a pressure relief valve. The probe consisted of three tubes: the first tube to apply static pressure (Ps) to the ear canal, the second tube to deliver sound to the external ear canal via an earphone, and the third tube to measure sound pressure in the external ear canal using a microphone. A specially designed cuff suitable for testing neonates was attached to the tip of the probe to obtain a hermetic seal during testing. This new SFI unit was controlled using LabView under Microsoft Windows. The SFI unit also performs HFT on infants as part of the test procedures.

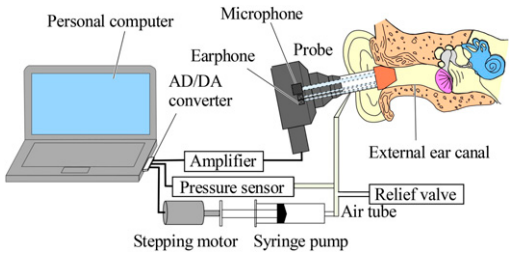

Figure 2. Block diagram of SFI meter used to test newborns in this study. The SFI meter consists of a personal computer, an analog-to-digital/digital-to-analog converter, a probe system, a stepping motor, a syringe pump, a pressure sensor, and a relief valve. This new SFI meter is controlled using LabView (Wada et al, 2016). (This figure appears in color in the online version of this article.)

During the SFI test, the SPL in the ear canal was measured as the frequency of the pure-tone stimulus was swept from 100 to 2200 Hz while the external auditory canal Ps was held constant at +200 daPa. This measurement was repeated with Ps reduced in 50 daPa steps down to −200 daPa. The entire SFI procedure was automated, taking <1 min to complete the test in each ear. A sweeping probe tone was delivered to the ear at 75 dB SPL to reduce the risk of eliciting an acoustic stapedial reflex. While the SFI results measured at multiple static pressures provide a comprehensive view of the acoustic-mechanical properties of the outer and middle ear, Murakoshi et al (2013) and Aithal et al (2014) found that measurements made at ambient pressure (0 daPa) can provide adequate clinical information about the status of the outer and middle ear. For the purpose of the present study, only measurements performed at ambient pressure were included in the analyses.

The pass/refer criteria for the SFI measures in the present study were based on normative data developed by Aithal et al (2014) using the same SFI unit. Aithal et al (2014) noted two regions of resonance in newborns, with the mean RF1 for the first resonance occurring at 287 Hz (90% range: 209–420 Hz) and the mean RF2 for the second resonance occurring at 1236 Hz (90% range: 830–1518 Hz). The first and second resonances refer to the resonances of the ear canal and middle ear, respectively. The authors' subsequent study (Aithal et al, 2016) showed that absence of the second resonance was associated with middle ear dysfunction in newborns. The presence of the second resonance with RF2 value of between 830 and 1518 Hz was considered a pass (indicating normal middle ear function).

Reference Standards and Pass/Refer Classification

At present, there is no complete agreement on which "reference" standard should be used to determine the test performance of diagnostic tests for the detection of disorders of the sound conduction pathways in newborns. Researchers have used both DPOAE (Margolis et al, 2003; Sanford et al, 2009; Hunter et al, 2010) and TEOAE (Kei et al, 2003; Vander Werff et al, 2007; Shahnaz, 2008) to determine the status of the middle ear such that absent or low level OAEs are suggestive of middle ear disorders in the absence of SN hearing loss that is secondary to cochlear damage. Although Norton et al (2000b) found no difference in the performance of TEOAE and DPOAE to detect hearing loss in newborns, the inclusion of both TEOAE and DPOAE as reference standards in the present study would be useful because the mechanism involved in generating OAEs by the two procedures are different. Furthermore, the two procedures demonstrate different susceptibility to background and biologic noise, resulting in different test outcomes (Rhoades et al, 1998; Shi et al, 2000; Norton et al, 2000a,b,c).

The present study used nine reference standards (four single tests and five test batteries) for determining the test performance of SFI. Table 1 shows the nine reference standards adopted in this study. While a single test such as DPOAE is useful in identifying a problem in the auditory pathway, it also has limitations that compromise its predictive accuracy because the DPOAE test alone does not distinguish conductive pathology from hair cell loss in the cochlea. For this reason, test battery reference standards involving a combination of tests were used in the present study. From a clinical perspective, newborns who passed a battery of tests involving HFT and DPOAE were more likely to have a normal sound conduction pathway (outer and middle ear) than those who passed DPOAE or HFT only (Aithal et al, 2013; 2015a). In the case of test battery reference standards, a strict test protocol was used (Keefe et al, 2003). With this protocol, the ear with a pass in all the tests in any given test battery was included in the "pass" group for that reference standard. For instance, with the HFT + DPOAE reference standard, only ears with a pass in both the HFT and DPOAE tests were included in the pass group for that test battery. Likewise, ears with a refer in each test of the test battery were included in the "refer" group for that reference standard. For example, with the HFT + TEOAE reference standard, only the ears with a refer in both the HFT and TEOAE tests were included in the "refer" group for that reference standard. While this strict test protocol provides clear separation between the pass and refer groups (Keefe et al, 2003; Aithal et al, 2015a), it excludes ears that have passed one test, but failed in the other test. If the ears that failed either of the tests are omitted, the test performance (sensitivity and specificity values) would be inflated. However, if the ears that failed in either the OAE or HFT test were classified in the "refer" category, the sensitivity and specificity values would be deflated. The present study, having adopted a strict test protocol, acknowledged this as a limitation of the study.

Table 1. Reference Standard Adopted in This Study Showing Number of Ears That Passed or Were Referred in Each Group

Reference Standard	Ears Passed	Ears Referred	Total
SFI	222	41	263
AABR	254	9	263
HFT	218	45	263
DPOAE	223	40	263
TEOAE	201	62	263
HFT + DPOAE	199	21	220
HFT + TEOAE	184	28	212
DPOAE + TEOAE	198	37	235
AABR + DPOAE	222	8	230
AABR + TEOAE	200	8	208

Table 1 shows the number of ears that passed or referred in each of the nine reference standards adopted in this study. For example, for the DPOAE reference standard, 40 ears were referred and 223 ears passed of 263 ears. When DPOAE was combined with HFT in a test battery (HFT + DPOAE) reference standard, the number of ears referred was reduced to 21 of a total of 220 ears. This decreased referral rate for the test battery reference standard was due to the use of a strict test protocol in which the outcomes were mixed across reference standards.

Data Analysis

All analyses were performed using the SPSS software (version 22; IBM, Armonk, NY). In the present study, RF2 values between 830 and 1518 Hz were considered a pass (normal) while RF2 values greater than 1518 Hz or absence were considered a refer (abnormal). This normative range was based on the results of a previous study (Aithal et al, 2014). The test performance of SFI was determined in terms of the sensitivity, specificity, and AROC.

The test performance of SFI could also be determined using likelihood ratio (LR) analysis. The LR of a test refers to improvement in the likelihood of making a correct diagnosis or identifying a condition. Positive LR (LR+) refers to improvement of the likelihood of correctly identifying the presence of a condition, whereas negative LR (LR−) refers to improvement of the likelihood of correctly identifying the absence of a condition. In general, a test with LR+ greater than 10 and LR− less than 0.1 is considered to be an effective test.

RESULTS

Table 2 shows the details of newborns included in the study. All newborns had an uneventful birth history with no risk factors for hearing loss (JCIH, 2007). The study included 263 ears (133 right and 130 left) from 188 healthy newborns (98 males and 90 females). Table 2 also shows the mean and standard deviation (SD) for gestational age (in weeks), birth weight (in grams), and age of testing (in hours) for 188 newborns. In order to maximize the data, each ear was considered independent of each other because each ear could have provided a different outcome (pass or refer). Although this maximized the available data, the authors are aware that there may be correlations in measurements between the right and left ears. However, a previous SFI normative study showed that there were no significant differences or interactions between ears and genders (Aithal et al, 2014).

An AROC value of 1 would indicate perfect agreement in pass/refer outcomes for the SFI and the comparison or test battery reference standard, whereas an AROC of 0.5 would indicate no better than chance performance. An AROC was computed to determine the test performance of SFI against four single test and five test battery reference standards. The results for the nine reference standards adopted in this study are shown in Table 3. Among the four single test reference standards, AROC was the greatest for HFT (0.82; 95% CI: 0.74–0.89) and smallest for AABR (0.61; 95% CI: 0.41–0.81). In comparison, among the five test battery reference standards, AROC was the greatest for the HFT+DPOAE reference standard (0.87; 95% CI: 0.78–0.96). AROC was the smallest for the AABR + TEOAE reference standard (0.66; CI: 0.45–0.88). AROC was significantly greater than 0.5 for the HFT, DPOAE, and TEOAE single test and for the HFT + DPOAE, HFT + TEOAE, and DPOAE + TEOAE test battery reference standards, as determined using the statistical procedure described by Hanley and McNeil (1982) ($p < 0.05$). These results indicate that the ability of SFI in identifying conductive conditions was lower when the AABR, AABR + DPOAE, and AABR + TEOAE reference standards were used because the AROCs were not significantly different from 0.5.

Table 3 also shows the 95% CI for AROC when SFI outcomes were compared with the outcomes of the nine reference standards. The CIs for the AABR, AABR + DPOAE, and AABR + TEOAE reference standards were broader than the other six reference standards. Although the CIs for DPOAE, TEOAE, and DPOAE + TEOAE were better, they were still broader to that of the HFT, HFT + DPOAE, and HFT + TEOAE reference standards. Table 4 shows the sensitivity, specificity, LR+, and LR− for SFI against both single and test battery reference standards. Among the four single reference standards, SFI had the highest sensitivity and specificity of 76% and 88%, respectively, against HFT. Among the test battery reference standards, SFI had the highest sensitivity and specificity of 86% and 88%, respectively, against HFT + DPOAE (Table 4).

As mentioned above, CIs for HFT, HFT + DPOAE, and HFT + TEOAE were similar and narrower compared to other reference standards (Table 3). In order to determine whether the AROCs of HFT, HFT + DPOAE, and HFT + TEOAE reference standards were significantly different from each other, a statistical test as described by Hanley and McNeil (1982) was applied using an online vassarstats test (http://vassarstats.net/roc_comp.html). The results showed no significant difference in AROC between the HFT and HFT + DPOAE ($p = 0.44$), HFT and HFT + TEOAE ($p = 1.00$), and HFT + DPOAE and HFT + TEOAE reference standards ($p = 0.48$).

Table 2. Details of Infants Included in This Study

	Male	Female	Total
Number of neonates	98	90	188
Right ear	29	29	58
Left ear	27	28	55
Bilateral	42	33	75
Gest age (weeks)			
Mean	39.4	39.3	39.3
SD	1.29	1.1	1.2
Birth weight (g)			
Mean	3,537.9	3,386.6	3,465.4
SD	414.8	490.7	457.8
Age at time of testing (hr)			
Mean	44.4	47	45.6
SD	22.8	19	21

Note: N = 188 infants; total ears = 263 (right = 133; left = 130).

DISCUSSION

The present study evaluated the test performance of SFI in terms of its ability to identify conductive conditions in newborns against various single test and test battery reference standards. As a single noninvasive gold standard does not exist for diagnosing conductive conditions in neonates, nine audiological test reference standards were used for comparison with SFI in this study.

The present study showed that the test performance of SFI against the HFT reference standard was higher than that against any of the AABR, DPOAE, and TEOAE reference standards as revealed by their respective AROC values (Table 3). These results imply that the test outcomes of SFI compare more favorably with that of the HFT test than with the AABR, DPOAE, and TEOAE tests, indicating the ability of SFI to detect conductive conditions in neonates. Since HFT could identify conductive conditions in newborns with high accuracy (Baldwin, 2006) and HFT's performance was as good as that of WBR (Prieve et al, 2013a,b), it is reasonable

Table 3. Test Performance of SFI against Different Reference Standards as Determined by AROC

Reference Standard	AROC	95% CI	Standard Error	Significance
AABR	0.61	0.41–0.81	0.10	0.26
HFT	0.82	0.74–0.89	0.04	0.00*
DPOAE	0.69	0.59–0.79	0.05	0.00*
TEOAE	0.62	0.54–0.71	0.04	0.00*
HFT + DPOAE	0.87	0.78–0.96	0.05	0.00*
HFT + TEOAE	0.82	0.72–0.91	0.05	0.00*
DPOAE + TEOAE	0.69	0.59–0.79	0.05	0.00*
AABR + DPOAE	0.67	0.45–0.88	0.11	0.12
AABR + TEOAE	0.66	0.45–0.88	0.11	0.12

Note: Null hypothesis: true area = 0.5; *significantly different from 0.5 with $p < 0.05$.

to infer that SFI identifies conductive disorders in newborns with reasonably good accuracy. Further analysis comparing SFI results with HFT outcomes revealed a sensitivity of 76% and specificity of 88% (Table 4). The LR+ and LR− were 6 and 0.3, respectively. Although LR did not reach the preferred values of LR+ greater than 10 or LR− less than 0.1, these results suggest that SFI may be a diagnostically useful test (Ebell, 2016).

In contrast, the test performance of SFI against the AABR reference standard was the lowest among the four single test reference standards. The AROC for AABR was 0.61 which was not significantly different from 0.5, indicating that SFI could not predict the outcomes of AABR more than chance. Studies have shown that AABR is not sensitive to slight/mild conductive hearing losses (Stapells, 2000; 2011; Aithal et al, 2012). Furthermore, a refer result in AABR may indicate a significant SN hearing loss which will not be detected by SFI.

The test performance of SFI against TEOAE and DPOAE was low (Tables 3 and 4) but significantly greater than 0.5, indicating that the test performance of SFI against these reference standards was better than chance in identifying conductive conditions in newborns (Table 3). One of the reasons for TEOAE and DPOAE not being ideal reference standards for conductive loss is that they are intended to detect auditory disorders up to the inner ear. TEOAE and DPOAE tests are also limited in their ability to accurately assess conductive hearing loss <2 kHz where conductive disorders greatly reduce reverse transmission of emissions at these frequencies. Furthermore, OAE test results are affected by environmental and physiologic noises which may produce a refer outcome in a normally hearing child. Like AABR, a refer result in an OAE test may indicate a significant SN hearing loss which will not be detected by the SFI.

Table 4. Table Showing the Sensitivity, Specificity, LR+, and LR− of SFI for Different Reference Standards

Reference Standard	Sensitivity (Estimate in %)	Specificity (Estimate in %)	LR+	LR−
AABR	44	78	2	0.7
HFT	76	88	6	0.3
DPOAE	55	83	3	0.5
TEOAE	42	83	3	0.7
HFT + DPOAE	86	88	7	0.2
HFT + TEOAE	75	88	6	0.3
DPOAE + TEOAE	54	83	3	0.6
AABR + DPOAE	50	82	3	0.6
AABR + TEOAE	50	83	3	0.6

The test performance of SFI against the HFT + DPOAE test battery reference standard was high with an AROC of 0.87, which was significantly better than that of AABR, DPOAE, TEOAE, and DPOAE + TEOAE (Table 3). This increased test performance was expected because a combination of tests may often detect a disease condition with higher accuracy than that of an individual test alone (Baughman et al, 2008; Naaktgeboren et al, 2013). The sensitivity of SFI was higher when it was compared to HFT + DPOAE then when it was compared to HFT alone. For instance, inclusion of HFT in a test battery with DPOAE (HFT + DPOAE) increased sensitivity of SFI from 55% to 86%. In general, the accuracy of SFI was increased with the inclusion of HFT in the test battery reference standard and decreased with the inclusion of AABR in the test battery reference standard. The LR+ and LR− were ≥6 and ≤0.3, respectively, with the inclusion of HFT in the test battery reference standard (Table 4). Although LR did not reach the preferred values of LR+ greater than 10 or LR− less than 0.1, these results suggest that SFI can be considered a diagnostically useful test with a moderate effect and that the likelihood of making a correct diagnosis that matched the other tests used to infer true condition was enhanced when a more strict reference standard was used.

Although it is not possible to directly compare the present investigation with other studies which have used wideband acoustic immittance measures, it is noted that the test performance of SFI against the DPOAE reference standard was better than that previously reported. For example, Sanford et al (2009) reported an AROC of 0.86 (95% CI: 0.80–0.89) for WBA and an AROC of 0.75

(95% CI: 0.68–0.80) for HFT against the DPOAE reference standard on day one of screening. However, on day two of screening, the AROC was reduced to 0.67 (95% CI: 0.45–0.83) for WBA and 0.54 (95% CI: 0.36–0.71) for HFT. In comparison, the present study showed an AROC of 0.69 (95% CI: 0.59–0.79) for SFI against the DPOAE reference standard which is better than that obtained by the Sanford et al (2009) study on day two of screening. The results of the present study are also better than the highest AROC of 0.67 for WBA at 1.25 kHz against a DPOAE reference standard obtained by Aithal et al (2015a). However, Hunter et al (2010) reported higher AROC values than those of the present study with an AROC of 0.82 and 0.90 for WBR at 1 and 2 kHz, respectively. The age of testing could be one of the reasons for these differences as the majority of infants were tested within the first 24–48 hr after birth in the Hunter et al (2010) study, which is earlier than the present study.

The sensitivity and specificity of SFI against the DPOAE reference standard in identifying conductive conditions, as shown in the present study, were 55% and 83%, respectively (Table 4). These results are consistent with the findings of Margolis et al (2003) who reported a sensitivity of 50% and specificity of 91% for HFT against a DPOAE reference standard. Swanepoel et al (2007) reported slightly better sensitivity of 57% and specificity of 95% for HFT against a DPOAE reference standard. In summary, the performance of SFI against DPOAE was comparable to that of HFT against DPOAE in identifying conductive conditions in newborns.

Clinical Application

The high test performance of SFI against the test battery (HFT + DPOAE) and single test (HFT) reference standards suggests that SFI is a valid measure of the function of the middle ear in newborns. Hence, SFI may be employed in UNHS programs as an adjunct test to the AABR screen. Similarly, SFI may also be employed in UNHS programs as an adjunct test to the DPOAE screen. The clinical information provided by the SFI test may be useful for prioritizing newborns for further diagnostic testing. A neonate who failed the SFI test but passed the AABR test will receive follow-up assessments to determine if the conductive condition has been resolved. A neonate who passed the SFI test, but failed in the AABR test would require further diagnostic assessments to determine the degree and nature of the hearing deficit. In the worst scenario when a neonate has failed both the SFI and AABR tests, a referral for diagnostic audiology assessment along with referral to an otolaryngologist is recommended because of the possibility of middle ear dysfunction along with SN hearing loss. In addition, the test performance of SFI justifies its application as a diagnostic test in UNHS follow-up testing. However, further research is required to evaluate the predictive accuracy of SFI in other age groups and ears with different conductive disorders.

Limitations

One of the limitations of the present study was the use of a strict protocol for determination of pass or refer status when a test battery reference standard was used. This strict protocol excluded the ears that "passed" in one test but "referred" in the other test within the test battery reference standard. The adoption of this strict protocol not only reduced the sample size, consequently reducing the power of the statistical analyses, but it also would have inflated the AROC, sensitivity, and specificity values for the SFI against the test battery reference standards.

Although it was relatively easy to test well-settled newborns, the SFI test took ~1 min to conduct in each ear. Since all tests had to be completed using different equipment in the same session, the removal and reinsertion of probes would sometimes disturb the neonate. Calming the neonate often increased the test duration. If the neonate became unsettled, the chance of having incomplete data collection increased. Further research is recommended using equipment that allows all tests to be done using a single probe. Further improvement in SFI instrumentation is needed to speed up the test if it is to be used as an assessment tool in newborn screening and diagnostics.

The outcome of the study could have also been influenced by the pass and refer criteria of some tests. For instance, the pass criterion for HFT was a single positive peak with positive peak admittance (Ya) of ≥0.2 mmho, while double or multiple peaks were considered a refer in this study. Similarly, the TEOAE criterion of at least a 3 dB SNR in three frequency regions (2, 3, and 4 kHz) and DPOAE criterion of minimum SNR of 6 dB and DPOAE amplitude greater than −6 dB SPL in at least three of four f2 frequencies (2, 3, 4, and 6 kHz) might not give optimal results. Furthermore, these pass criteria did not adequately assess frequencies <2 kHz where the impact of middle ear disorders on audition is more prominent.

The test performance of SFI may vary depending on the time of screening during the postnatal period. The mean age of screening newborns in the present study was 45.6 hr, but with a substantial SD of 21 hr. Studies have shown that transient conductive conditions due to the presence of vernix and/or mesenchyme in the outer and middle ear may occur within the first 48 hr of birth (Doyle et al, 2000; Allen et al, 2005; Sanford et al, 2009). The transient conductive conditions would affect the referral rates of neonates in all screening tests which, in turn, affect the test performance of SFI. Hence, it is very important to consider the time of screening after delivery

as a contributing factor when comparing the test performance of different protocols.

There is lack of a "true" gold standard for testing middle ear function in newborns. Although test battery reference standards were used to determine normal outer/middle ear sound conduction function, a pass in all of these tests cannot definitively rule out slight outer/middle ear dysfunction in newborns (Aithal et al, 2012).

CONCLUSIONS

The test performance of SFI was compared against four single test and five test battery reference standards in this study. The test performance of SFI against HFT with an AROC of 0.82 and against HFT + DPOAE with an AROC of 0.87 indicates that SFI can accurately identify conductive conditions in newborns. Hence, the SFI test may be useful for both screening and diagnostic assessments in newborns.

Acknowledgments. The authors are thankful to Associate Professor Andrew Swanston, Director ENT Department, Townsville Hospital and Health Service; and Katrina Roberts, Nursing Director, Health and Wellbeing Service Group, Townville Hospital and Health Service for their support. The authors are also thankful to Marissa Edmondson, Jewelie-Ann Wright, Rowena Lyons, Nicky Audas, and Jackie Bunt from the Health and Wellbeing Service Group; Sreedevi Aithal and Karen Nielsen from the Audiology Department for their assistance in data collection; and Ben Mason, Biomedical Technology Service for calibration and maintenance of equipment. The authors gratefully acknowledge the two anonymous reviewers for their valued comments on this manuscript.

REFERENCES

Aithal S, Aithal V, Kei J, Driscoll C. (2012) Conductive hearing loss and middle ear pathology in young infants referred through a newborn universal hearing screening program in Australia. *J Am Acad Audiol* 23(9):673–685.

Aithal S, Kei J, Driscoll C, Khan A. (2013) Normative wideband reflectance measures in healthy neonates. *Int J Pediatr Otorhinolaryngol* 77(1):29–35.

Aithal S, Kei J, Driscoll C, Khan A, Swanston A. (2015a) Wideband absorbance outcomes in newborns: a comparison with high-frequency tympanometry, automated brainstem response, and transient evoked and distortion product otoacoustic emissions. *Ear Hear* 36(5):e237–e250.

Aithal V, Kei J, Driscoll C, Murakoshi M, Wada H. (2016) Effects of ear canal static pressure on the dynamic behaviour of outer and middle ear in newborns. *Int J Pediatr Otorhinolaryngol* 82:64–72.

Aithal V, Kei J, Driscoll C, Swanston A, Murakoshi M, Wada H. (2015b) Sweep frequency impedance measures in Australian Aboriginal and Caucasian neonates. *Int J Pediatr Otorhinolaryngol* 79(7):1024–1029.

Aithal V, Kei J, Driscoll C, Swanston A, Roberts K, Murakoshi M, Wada H. (2014) Normative sweep frequency impedance measures in healthy neonates. *J Am Acad Audiol* 25(4):343–354.

Alaerts J, Luts H, Wouters J. (2007) Evaluation of middle ear function in young children: clinical guidelines for the use of 226- and 1,000-Hz tympanometry. *Otol Neurotol* 28(6):727–732.

Allen JB, Jeng PS, Levitt H. (2005) Evaluation of human middle ear function via an acoustic power assessment. *J Rehabil Res Dev* 42(4, Suppl 2):63–78.

Baldwin M. (2006) Choice of probe tone and classification of trace patterns in tympanometry undertaken in early infancy. *Int J Audiol* 45(7):417–427.

Balkany TJ, Berman SA, Simmons MA, Jafek BW. (1978) Middle ear effusions in neonates. *Laryngoscope* 88(3):398–405.

Baughman AL, Bisgard KM, Cortese MM, Thompson WW, Sanden GN, Strebel PM. (2008) Utility of composite reference standards and latent class analysis in evaluating the clinical accuracy of diagnostic tests for pertussis. *Clin Vaccine Immunol* 15(1):106–114.

Boone RT, Bower CM, Martin PF. (2005) Failed newborn hearing screens as presentation for otitis media with effusion in the newborn population. *Int J Pediatr Otorhinolaryngol* 69(3):393–397.

Boudewyns A, Declau F, Van den Ende J, Van Kerschaver E, Dirckx S, Hofkens-Van den Brandt A, Van de Heyning P. (2011) Otitis media with effusion: an underestimated cause of hearing loss in infants. *Otol Neurotol* 32(5):799–804.

Buch NH, Jorgensen MB. (1964) Embroynic connective tissue in the tympanic cavity of the fetus and newborn. *Acta Otolaryngica.* 58:111–126.

Clemens CJ, Davis SA. (2001) Minimizing false-positives in universal newborn hearing screening: a simple solution. *Pediatrics* 107(3):E29.

Clemens CJ, Davis SA, Bailey AR. (2000) The false-positive in universal newborn hearing screening. *Pediatrics* 106(1):E7.

Derkay CS, Bluestone CD, Thompson AE, Stephenson JS, Kardtzke D. (1988) Otitis media in the pediatric intensive care unit: a prospective study of its incidence, prevalence, bacteriology, etiology and treatment. *Ann Otol Rhinol Laryngol* 97(Suppl 133):56:57.

Doyle KJ, Burggaaff B, Fujikawa S, Kim J, Macarthur CJ. (1997) Neonatal hearing screening with otoscopy, auditory brain stem response and otoacoustic emissions. *Otolaryngol Head Neck Surg* 116:597–603.

Doyle KJ, Kong YY, Strobel K, Dallaire P, Ray RM. (2004) Neonatal middle ear effusion predicts chronic otitis media with effusion. *Otol Neurotol* 25(3):318–322.

Doyle KJ, Rodgers P, Fujikawa S, Newman E. (2000) External and middle ear effects on infant hearing screening test results. *Otolaryngol Head Neck Surg* 122(4):477–481.

Ebell MH. (2016) Evidence-based practice for health professions: a free online course. Chapter 4: Diagnostic Tests-ROC Curves. Athens, Georgia: The University of Georgia, College of Public Health.

Gravel JS, White KR, Johnson JL, Widen JE, Vohr BR, James M, Kennalley T, Maxon AB, Spivak L, Sullivan-Mahoney M, Weirather Y, Meyer S. (2005) A multisite study to examine the efficacy of the otoacoustic emission/automated auditory brainstem response newborn hearing screening protocol: recommendations for policy, practice, and research. *Am J Audiol* 14(2):S217–S228.

Hamanishi S, Murakoshi M, Steele C, Wada H, Puria S. (2015) A hybrid ear canal and middle ear model approach for neonates. 38th Annual Winter Meeting of Association for Research in Otolaryngology. Washington, DC February 21–25, 2015

Hanley JA, McNeil BJ. (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 143(1):29–36.

Holster IL, Hoeve LJ, Wieringa MH, Willis-Lorrier RM, de Gier HHW. (2009) Evaluation of hearing loss after failed neonatal hearing screening. *J Pediatr* 155(5):646–650.

Holte L, Cavanaugh RM, Jr, Margolis RH. (1990) Ear canal wall mobility and tympanometric shape in young infants. *J Pediatr* 117(1 Pt 1):77–80.

Hunter LL, Blankenship CM. (2017) Middle ear measurement. In: Tharpe AM, Seewald R, eds. *Comprehensive Handbook of Pediatric Audiology*. San Diego, CA: Plural Publishing.

Hunter LL, Feeney MP, Lapsley Miller JA, Jeng PS, Bohning S. (2010) Wideband reflectance in newborns: normative regions and relationship to hearing-screening results. *Ear Hear* 31(5):599–610.

Hunter LL, Keefe DH, Feeney MP, Fitzpatrick DF, Lin L. (2015) Longitudinal development of wideband reflectance tympanometry in normal and at-risk infants. *Hear Res* 340:3–14.

Joint Committee on infant Hearing (JCIH). (2007) Year 2007 position statement: principles and guidelines for early hearing detection and intervention programs. *Pediatrics* 120:898–921.

Keefe DH, Bulen JC, Arehart KH, Burns EM. (1993) Ear-canal impedance and reflection coefficient in human infants and adults. *J Acoust Soc Am* 94(5):2617–2638.

Keefe DH, Feeney MP. (2009) Principles of acoustic imittance and acoustic transfer functions. In: Katz J, Burkhard RF, Medwetsky L, eds. *Handbook of Clinical Audiology*. Baltimore, MD: Lippincott, Williams and Wilkins.

Keefe DH, Folsom RC, Gorga MP, Vohr BR, Bulen JC, Norton SJ. (2000) Identification of neonatal hearing impairment: ear-canal measurements of acoustic admittance and reflectance in neonates. *Ear Hear* 21(5):443–461.

Keefe DH, Gorga MP, Neely ST, Zhao F, Vohr BR. (2003) Ear-canal acoustic admittance and reflectance measurements in human neonates. II. Predictions of middle-ear in dysfunction and sensorineural hearing loss. *J Acoust Soc Am* 113(1):407–422.

Keefe DH, Levi E. (1996) Maturation of the middle and external ears: acoustic power-based responses and reflectance tympanometry. *Ear Hear* 17(5):361–373.

Kei J, Allison-Levick J, Dockray J, Harrys R, Kirkegard C, Wong J, Maurer M, Hegarty J, Young J, Tudehope D. (2003) High-frequency (1000 Hz) tympanometry in normal neonates. *J Am Acad Audiol* 14(1):20–28.

Kei J, Mazlan R. (2012) High-frequency (1000 Hz) tympanometry: clinical applications. In: Kei J, Zhao F, eds. *Assessing Middle Ear Function in Infants*. San Diego, CA: Plural Publishing, 39–67.

Kei J, Mazlan R, Hickson L, Gavranich J, Linning R. (2007) Measuring middle ear admittance in newborns using 1000 Hz tympanometry: a comparison of methodologies. *J Am Acad Audiol* 18(9): 739–748.

Kei J, Zhao F. (2012) The challenge of assessing middle ear function in young infants: summary and future directions. In: Kei J, Zhao F, eds. *Assessing Middle Ear Function in Infants*. San Diego, CA: Plural Publishing, 163–177.

Kok MR, van Zanten GA, Brocaar MP. (1992) Growth of evoked otoacoustic emissions during the first days postpartum: a preliminary report. *Audiology* 31(3):140–149.

Margolis RH, Bass-Ringdahl S, Hanks WD, Holte L, Zapala DA. (2003) Tympanometry in newborn infants—1 kHz norms. *J Am Acad Audiol* 14(7):383–392.

Margolis RH, Shanks JE. (1990) Tympanometry: basic principles and clinical applications. In: Rintelmann WF, ed. *Hearing Assessment*. Austin, TX: Pro-Ed, 179–246.

Mason JA, Herrmann KR. (1998) Universal infant hearing screening by automated auditory brainstem response measurement. *Pediatrics* 101(2):221–228.

Mazlan R, Kei J, Hickson L. (2009) Test-retest reliability of the acoustic stapedial reflex test in healthy neonates. *Ear Hear* 30(3):295–301.

Mehl AL, Thomson V. (1998) Newborn hearing screening: the great omission. *Pediatrics* 101:E4.

Mehl AL, Thomson V. (2002) The Colorado newborn hearing screening project, 1992-1999: on the threshold of effective population-based universal newborn hearing screening. *Pediatrics* 109:E7.

Merchant GR, Horton NJ, Voss SE. (2010) Normative reflectance and transmittance measurements on healthy newborn and 1-month-old infants. *Ear Hear* 31(6):746–754.

Murakoshi M, Yoshida N, Sugaya M, Ogawa Y, Hamanishi S, Kiyokawa H, Kakuta R, Yamada M, Takahashi R, Tanigawara S, Matsutani S, Kobayashi T, Wada H. (2013) Dynamic characteristics of the middle ear in neonates. *Int J Pediatr Otorhinolaryngol* 77(4): 504–512.

Murakoshi M, Zhao F, Wada H. (2012) Current developments in the clinical application of the sweep frequency impedance (SFI) in assessing middle ear dysfunction. In: Kei J, Zhao F, eds. *Assessing Middle Ear Function in Infants*. San Diego, CA: Plural Publishing, 107–129.

Naaktgeboren CA, Bertens LC, van Smeden M, de Groot JA, Moons KG, Reitsma JB. (2013) Value of composite reference standards in diagnostic research. *BMJ* 347:f5605.

Norton SJ, Gorga MP, Widen JE, Folsom RC, Sininger Y, Cone-Wesson B, Vohr BR, Fletcher KA. (2000a) Identification of neonatal hearing impairment: summary and recommendations. *Ear Hear* 21(5):529–535.

Norton SJ, Gorga MP, Widen JE, Folsom RC, Sininger Y, Cone-Wesson B, Vohr BR, Mascher K, Fletcher K. (2000b) Identification of neonatal hearing impairment: evaluation of transient evoked otoacoustic emission, distortion product otoacoustic emission, and auditory brain stem response test performance. *Ear Hear* 21(5):508–528.

Norton SJ, Gorga MP, Widen JE, Folsom RC, Sininger Y, Cone-Wesson B, Vohr BR, Fletcher KA. (2000c) Identification of neonatal hearing impairment: a multicenter investigation. *Ear Hear* 21(5):348–356.

Paradise JL. (1976) Pediatrician's view of middle ear effusions: more questions than answers. *Ann Otol Rhinol Laryngol* 85(2, Suppl 25 Pt 2):20–24.

Paradise JL. (1980) Otitis media in infants and children. *Pediatrics* 65(5):917–943.

Paradise JL. (1981) Otitis media during early life: how hazardous to development? A critical review of the evidence. *Pediatrics* 68(6): 869–873.

Prieve BA, Feeney MP, Stenfelt S, Shahnaz N. (2013a) Prediction of conductive hearing loss using wideband acoustic immittance. *Ear Hear* 34(1, Suppl):54S–59S.

Prieve BA, Vander Werff KR, Preston JL, Georgantas L. (2013b) Identification of conductive hearing loss in young infants using tympanometry and wideband reflectance. *Ear Hear* 34(2): 168–178.

Purdy SC, Williams SM. (2002) High frequency tympanometry: a valid and reliable immittance test protocol for young infants. *NZ Audiolog Soc Bull* 10:9–24.

Qi L, Liu H, Lutfy J, Funnell WR, Daniel SJ. (2006) A nonlinear finite-element model of the newborn ear canal. *J Acoust Soc Am* 120(6):3789–3798.

Rhodes MC, Margolis RH, Hirsch JE, Napp AP. (1999) Hearing screening in the newborn intensive care nursery: comparison of methods. *Otolaryngol Head Neck Surg* 120(6):799–808.

Rhoades K, McPherson B, Smyth V, Kei J, Baglioni A. (1998) Effects of background noise on click-evoked otoacoustic emissions. *Ear Hear* 19(6):450–462.

Rosenfeld RM, Culpepper L, Doyle KJ, Grundfast KM, Hoberman A, Kenna MA, Lieberthal AS, Mahoney M, Wahl RA, Woods CR Jr, Yawn B; American Academy of Pediatrics Subcommittee on Otitis Media with Effusion; American Academy of Family Physicians; American Academy of Otolaryngology–Head and Neck Surgery. (2004) Clinical practice guideline: otitis media with effusion. *Otolaryngol Head Neck Surg* 130(5, Suppl): S95–S118.

Ruah CB, Schachern PA, Zelterman D, Paparella MM, Yoon TH. (1991) Age-related morphologic changes in the human tympanic membrane. A light and electron microscopic study. *Arch Otolaryngol Head Neck Surg* 117(6):627–634.

Sanford CA, Keefe DH, Liu Y-W, Fitzpatrick D, McCreery RW, Lewis DE, Gorga MP. (2009) Sound-conduction effects on distortion-product otoacoustic emission screening outcomes in newborn infants: test performance of wideband acoustic transfer functions and 1-kHz tympanometry. *Ear Hear* 30(6):635–652.

Sanford CA, Schooling T, Frymark T. (2012) Determining the presence or absence of middle ear disorders: an evidence-based systematic review on the diagnostic accuracy of selected assessment instruments. *Am J Audiol* 21(2):251–268.

Sangster L. (2011) *Critical Review: Can Wideband Energy Reflectance Be Used in Newborn Hearing Screening to Detect Transient Middle Ear Dysfunction and to Interpret Screening Results?* Ontario, Canada: University of Western Ontario, 1–6.

Saunders JC, Kaltenback JA, Relkin EM. (1983) The structural and functional development of the outer and middle ear. In: Romand R, Romand MR, eds. *Development of Auditory and Vestibular Systems*. New York, NY: Academic Press, 3–25.

Shahnaz N. (2008) Wideband reflectance in neonatal intensive care units. *J Am Acad Audiol* 19(5):419–429.

Shi S, Kei J, Murdoch B, McPherson B, Smyth V, Latham S, Loscher J. (2000) Paediatric hearing screening in the community: a comparison of outcomes from transient evoked and distortion product otoacoustic emission measures. *Scand Audiol* 29(2): 83–92.

Shurin PA, Pelton SI, Klein JO. (1976) Otitis media in the newborn infant. *Ann Otol Rhinol Laryngol* 85(2, Suppl 25 Pt 2): 216–222.

Silverman CA. (2010) Multifrequency tympanometry in infants and children. *Audiologyonline*, www.audiologyonline.com/articles/multifrequency-tympanometry-in-infants-and-852.

Stapells DR. (2000) Threshold estimation by the tone evoked auditory brainstem response: a literature meta-analysis. *J Speech Lang Pathol Audiol* 24:74–83.

Stapells DR. (2011) Frequency-specific threshold assessment in young infants using the transient ABR and the brainstem ASSR. In: Seewald R, Tharpe AM, eds. *Comprehensive Handbook of Pediatric Audiology*. San Diego, CA: Plural Publishing, 409–448.

Swanepoel W, Werner S, Hugo R, Louw B, Owen R, Swanepoel A. (2007) High frequency immittance for neonates: a normative study. *Acta Otolaryngol* 127(1):49–56.

Vander Werff KR, Prieve BA, Georgantas LM. (2007) Test-retest reliability of wideband reflectance measures in infants under screening and diagnostic test conditions. *Ear Hear* 28(5):669–681.

Vohr BR, Carty LM, Moore PE, Letourneau K. (1998) The Rhode Island Hearing Assessment Program: experience with statewide hearing screening (1993-1996). *J Pediatr* 133(3):353–357.

Wada H, Kobayashi T. (1990) Dynamical behavior of middle ear: theoretical study corresponding to measurement results obtained by a newly developed measuring apparatus. *J Acoust Soc Am* 87(1):237–245.

Wada H, Koike T, Kobayashi T. (1998) Clinical applicability of the sweep frequency measuring apparatus for diagnosis of middle ear diseases. *Ear Hear* 19(3):240–249.

Wada H, Murakoshi M, Aithal V, Swanston A. (2016) Dynamic characteristic of the outer and middle ears in neonates. 33rd World Congress of Audiology, Vancouver, Canada.

Wada H, Ohyama K, Kobayashi T, Koike T, Noguchi S. (1995) Effect of middle ear on otoacoustic emissions. *Audiology* 34(4): 161–176.

Zhao F, Wang J. (2012) Assessing middle ear function in humans using multifrequency tympanometry: an overview. In: Kei J, Zhao F, eds. *Assessing Middle Ear Function in Infants*. San Diego, CA: Plural Publishing, 87–106.