Supplement

A Multisite Study to Examine the Efficacy of the Otoacoustic Emission/Automated Auditory Brainstem Response Newborn Hearing Screening Protocol: Results of Visual Reinforcement Audiometry

Judith E. Widen
University of Kansas Medical Center, Kansas City

Jean L. Johnson
University of Hawaiì, Honolulu

Karl R. White
National Center for Hearing Assessment and Management, Utah State University, Logan

Judith S. Gravel
Albert Einstein College of Medicine and Jacobi Medical Center, Bronx, NY

Betty R. Vohr
Women and Infants Hospital, Providence, RI

Michele James
Arnold Palmer Hospital for Children and Women, Orlando, FL

Teresa Kennalley
Via Christi Regional Medical Center, Wichita, KS

Antonia B. Maxon
New England Center for Hearing Rehabilitation, Hampton, CT

Lynn Spivak
Long Island Jewish Medical Center, New Hyde Park, NY

Maureen Sullivan-Mahoney
Good Samaritan Hospital, Cincinnati, OH

Yusnita Weirather
Kapiòlani Medical Center for Women and Infants, Honolulu, HI

Sally Meyer
National Center for Hearing Assessment and Management, Utah State University, Logan

American Journal of Audiology • Vol. 14 • S200–S216 • December 2005 • © American Speech-Language-Hearing Association
1059-0889/05/1402-S200

Purpose: This 3rd of 4 articles on a study of the efficacy of the 2-stage otoacoustic emission/automated auditory brainstem response (OAE/A-ABR) newborn hearing screening protocol describes (a) the behavioral audiometric protocol used to validate hearing status at 8–12 months of age, (b) the hearing status of the sample, and (c) the success of the visual reinforcement audiometry (VRA) protocol across 7 sites.
Method: A total of 973 infants who failed OAE but passed A-ABR, in one or both ears, during newborn screening were tested with a VRA protocol, supplemented by tympanometry and OAE screening at age 8–12 months.
Results: VRA audiograms (1.0, 2.0, and 4.0 kHz) were obtained for 1,184 (82.7%) of the 1,432 study ears. Hearing loss was ruled out in another 100 ears by VRA in combination with OAE, for a total of 88.7% of the study sample. Permanent hearing loss was identified in 30 ears of 21 infants. Sites differed in their success with the VRA protocol.
Conclusions: Continued monitoring of hearing beyond the newborn period is an important component of early detection of hearing loss. Using a structured protocol, VRA is an appropriate test method for most, but not all, infants. A battery of test procedures is often needed to adequately delineate hearing loss in infants. Examiner experience appears to be a factor in successful VRA.

Key Words: visual reinforcement audiometry, otoacoustic emissions, automated brainstem response audiometry, tympanometry

This is the third article in a series of four that describe the results of a multisite study of the efficacy of the two-stage otoacoustic emission/automated auditory brainstem response (OAE/A-ABR) newborn hearing screening protocol. The two-stage protocol is one in which all infants are screened first with OAEs. No additional testing is done for those who pass the OAE screening. Infants who fail the OAE screen are then screened with A-ABR to determine whether a referral for further diagnostic testing is needed. One of the reasons that many hospitals have adopted this protocol is to reduce the referral rate that has been reported for programs that screen with OAEs alone (Gravel et al., 2000; Vohr et al., 2001). The Centers for Disease Control and Prevention funded the present study to determine whether infants with congenital permanent hearing loss (PHL) were missed with the two-stage protocol. Full details of the background and rationale for the study are described in Johnson, White, Widen, Gravel, Vohr, et al. (2005). To answer the question of how many infants who fail the OAE and pass the A-ABR have PHL at 8–12 months of age, 86,634 infants were screened at seven birthing centers using a two-stage OAE/A-ABR hearing screening protocol. Of the infants who failed the OAE but passed the A-ABR, 1,524 were enrolled in the study and asked to return for diagnostic audiologic evaluation at age 8 months. Of those, 973 infants (63.4% of the study sample) returned for the follow-up evaluation. The study design and results of the overall research question are presented by White et al. (2005). The present article focuses specifically on the visual reinforcement audiometry (VRA) protocol used to determine hearing status and on the results that were obtained for the 973 infants (1,432 ears) who returned for diagnostic evaluation when they were 8–12 months old. In the fourth and final article in the series, Gravel et al. (2005) consider the issues and implications of the study's findings.

Although hearing is a behavioral response, identification of hearing loss in newborns is not possible using a behavioral technique. Objective, physiological measures such as ABR and OAEs have been found to be a reliable indictor of hearing loss in newborns. However, to validate the presence or absence of a hearing loss, a behavioral measure of hearing is required. Previous studies have shown that valid behavioral methods can accurately delineate hearing loss in infants once they have reached a developmental age of about 6 months (cf. Day, Bamford, Parry, Shepherd, & Quigley, 2000; Gravel & Wallace, 2000; Sabo, Paradise, Kurs-Lasky, & Smith, 2003; Talbott, 1987; Wilson & Thompson, 1984). Over the past 2 decades, a few studies have attempted to validate newborn tests with later behavioral testing (Lutman, Davis, Fortnum, & Wood, 1997; Mason, Davis, Wood, & Farnsworth, 1998; Shannon, Felix, Krumholz, Goldstein, & Harris, 1984; Swigonski, Shallop, Bull, & Lemons, 1987; van Straaten, Groote, & Oudesluys-Murphy, 1996; Watson, McClelland, & Adams, 1996). In none of these studies were ear-specific pure-tone behavioral audiograms attempted.

For the National Institute on Deafness and Other Communication Disorders (NIDCD) study of identification of neonatal hearing impairment (INHI), Norton et al. (2000) employed a VRA protocol using insert earphones to validate the earlier newborn measures. The INHI study was a multisite clinical trial that examined the three physiological measures used most frequently for newborn hearing screening—transient (click) evoked otoacoustic emissions (TEOAEs), distortion product otoacoustic emissions (DPOAEs), and ABR. The purpose of that study was to determine the sensitivity and specificity of the three measures in detecting congenital hearing loss. A total of 4,911 babies who were "at risk" (i.e., babies from the neonatal intensive care units [NICUs] and well babies with one or more of the Joint Committee on Infant Hearing, 1994, risk factors for hearing loss) were followed and evaluated using VRA at 8–12 months corrected age. In that study, 94% of infants who returned for a diagnostic

evaluation at 8–12 months old (64% of the original sample) were successfully tested with VRA using insert earphones (Widen et al., 2000).

To address the question of whether the two-stage OAE/A-ABR protocol was missing infants with a hearing loss, the investigators of the present study chose to evaluate the status of the infants' hearing using a VRA protocol adapted from the one used in the NIDCD/INHI study.

The three purposes of this article are to (a) describe the protocol used to determine hearing status at 8–12 months of age, (b) describe the results for the group as a whole and particularly for the infants/ears determined to have PHL, and (c) evaluate the success of the protocol and describe, via case reports, some of the challenges in implementing such a follow-up program and doing diagnostic audiologic assessments of 8–12-month-old babies.

Method

Participants

As described in White et al. (2005), the study sample was drawn from seven birthing centers across the United States, from families who represented the socioeconomic and ethnic diversity of the U.S. population. Ninety percent of the infants were recruited from well baby nurseries, 10% from NICUs. Of the 3,462 babies who were eligible (i.e., failed OAE, passed A-ABR for at least one ear), consent to participate was received from the parents of 1,524, and the babies were enrolled in the study. After the babies were 7 months adjusted age, all were asked to return for diagnostic assessment. A total of 973, or 63.8%, returned for at least one diagnostic session. These 973 study babies who actually returned for diagnostic evaluation are the focus of this article.

Test Sites

The diagnostic evaluations were conducted in seven audiology clinics associated with the hospitals that participated in the project: Good Samaritan Hospital in Cincinnati, OH; Kapiòlani Medical Center in Honolulu, HI; Jacobi Medical Center in the Bronx, NY; Arnold Palmer Medical Center in Orlando, FL; Via Christi Regional Medical Center in Wichita, KS; Women and Infants Hospital in Providence, RI; and the Long Island Jewish Speech and Hearing Center in New Hyde Park, NY.

Equipment

Evaluations at all seven audiology facilities were conducted in sound-treated test booths. Audiometers calibrated to American National Standards Institute (ANSI, 1996) standards for insert earphones (2-cc coupler) were used for testing. At six of the seven sites, the test room was arranged with four reinforcer toys (from Intelligent Hearing Systems), two on each side of the infant, employed to provide reinforcement for correct head turn responses. At one site, there were three reinforcer boxes, stacked to one side, and the audiometer and reinforcement system were controlled by one examiner who was seated in the test room with the parent and the child. Through the study, reinforcement equipment was provided to sites that needed an upgrade to meet the protocol requirements.

Equipment used for tympanometric measurements varied across sites and included the GSI 33 Middle Ear Analyzer, GSI Tympstar, and Madsen Zodiac 3000. All provided comparable measures of ear canal volume, peak-compensated static admittance, middle ear peak pressure, and tympanometric width using a 226-Hz probe tone. Either TEOAE or DPOAE measures were accepted because OAE equipment varied from site to site, and even within sites, and included the Biologic AuDx and Scout; GSI 60; Otodynamics EchoCheck, Echoport, and ILO88; and Starkey DP2000.

Personnel

The diagnostic evaluations were conducted by certified, licensed audiologists, all of whom had considerable experience testing infants and young children. At six sites, one examiner was located at the audiometer with control of the reinforcement system while another examiner was seated in the test room with the infant and parent. In most instances, this in-room examiner was an audiologist; in a few instances, the in-room examiner was an audiology student or assistant who had been trained as a VRA assistant. On rare occasions when a second examiner was not available, the parent was instructed to maintain the appropriate, attentive state for the infant's testing. As stated previously, one facility was equipped with an automated system that allowed a single examiner to maintain the infant's attention while also operating the equipment.

Training and practice on the protocol developed for this study were provided by the diagnostic evaluation coordinator (J. Widen), who traveled to each of the sites before the study infants reached 7 months and spent part of a day in practice sessions with the audiologists who would be conducting the diagnostic evaluations for the study infants. She and the audiologists also communicated throughout the data collection period regarding questions about implementing the protocol and reporting the data.

Procedures

At the time of enrollment in the study, the parents of all participants had been told that their newborn had passed the newborn screening based on the A-ABR result. When they were invited to enroll in the study, they were informed that they would be asked to return for a diagnostic evaluation when their baby was 7 months old. Regular communication was maintained with the families by postcards at 2-month intervals to ensure that the study stayed aware of changes in address and to increase the likelihood of participation (White et al., 2005). A small remuneration, generally $20, was provided to reimburse families for the costs of travel, parking, and child care, whenever they returned for an evaluation.

The optimum goal of the diagnostic evaluation was to obtain, in one visit, the following: (a) pure-tone minimum response levels (MRLs) of 15 dB HL at 0.5, 1.0, 2.0, and

4.0 kHz and a speech awareness threshold; (b) tympanometry; and (c) OAE screening (either DPOAEs or TEOAEs) for each ear. Recognizing that infants do not always cooperate with *optimum* goals, and that project funding could not cover multiple test sessions for all infants, a minimum goal was defined for the purposes of considering the data complete for analysis: MRLs at 1.0, 2.0, and 4.0 kHz in the study ear(s) with good confidence about their reliability and, if MRLs were elevated above 20 dB HL, no evidence of transient middle ear effusion per tympanometry. If these minimal goals could not be achieved during the first evaluation, a second evaluation was scheduled for a later time. In addition, realizing that the incidence of otitis media with effusion (OME) in this age group is high, it was expected that some infants would have transient hearing loss due to OME on the day of their evaluation and repeat visits would need to be scheduled for later, after the resolution of fluid.

The protocol was developed jointly by the coinvestigators of the study. The protocol for VRA was based on the protocol used in the NIDCD/INHI multicenter consortium study (Widen et al., 2000). Like that earlier protocol, importance was placed on ensuring that the infants' behavioral responses were under stimulus control, that is, that the strength of the discriminative stimulus (tones, speech) in producing the desired response (head turn) was maintained throughout the test session. Stimulus control was demonstrated with probe and control trials. Probe trials—stimulus presentations provided at suprathreshold levels—were used to confirm conditioning at the beginning of the test session and throughout the test session. Control trials—observation intervals in which no stimulus was present but the examiner noted head turn behavior and provided no reinforcement—were used to ensure that head turn responses were linked to stimulus presentation and not simply random movements.

A summary of the protocol is provided in Appendix A. Testing began at a level that was presumed to be audible to the majority of participants, in this case 35 dB HL. If the infant responded with a head turn, reinforcement was provided by activating one of the toys on the side to which the infant turned. A subsequent correct head turn during stimulus presentation at the same level resulted in reinforcement, and the test protocol was begun without training.

The infant was given two chances for an unconditioned response, one to the 35-dB stimulus, another at 55 dB HL. If the infant did not turn spontaneously toward the side stimulated, training (conditioning) trials were begun by pairing the reinforcer with the stimulus at 55 dB HL. After two paired trials, a probe trial was presented, with reinforcement provided only for a correct head turn response. A second correct probe trial was required before the search for MRL began. If the infant did not respond to this initial sequence of conditioning and probe trials, the training sequence of paired trials was repeated by first raising the signal level to 75 dB HL, then by changing the stimulus type (another frequency or speech), the test ear, or mode of presentation (usually to sound field presentation rather than earphone presentation). Conditioning was continued until the infant demonstrated stimulus control or until all options for training had been exhausted.

Control trials were inserted at a rate of 25%–30% but were actually printed at a rate of 50% on the data sheets (see Appendix B). This frequency rate was considered particularly important for the infants with normal hearing who could likely hear, and be rewarded for, every stimulus presented. If the infant responded during control trials more than 30% of the time, test validity was questioned and the infant was reconditioned after a break or rescheduled for another visit.

Tympanometric measures of ear canal volume, peak-compensated static admittance, middle ear peak pressure, and tympanometric width were recorded for each ear. Because OAE equipment varied across sites, examiners simply reported the results as the presence of an OAE at four frequency bands centered at 1.0, 2.0, 3.0, and 4.0 kHz, and whether the OAE signal-to-noise ratio (SNR) was <3 dB, 3–6 dB, or >6 dB.

Historical information about health status and ear infections since the newborn period was ascertained at each visit. Also reported were details about the test session such as examiners' names, their confidence in the results, type of stimulus (tones or speech) and transducer (earphones or speaker) used for beginning conditioning trials, length of test session (in number of trials and time), and a disposition (whether the follow-up protocol was completed, and if not, possible reasons why). The data collection form is shown in Appendix C.

Results

Completeness of Data Set

To answer the study question about whether babies with a PHL were missed using the two-stage (OAE/A-ABR) protocol, it was necessary to bring back as many of the infants as possible. Table 1 shows that of the 1,524 infants enrolled in the study, 973, or 63.8%, returned for diagnostic evaluation. Those 973 infants represented 1,432 ears that failed OAE but passed A-ABR screening in the hospital nursery. In addition, the other ear, which was not a "study ear" (an ear that passed the initial OAE), was evaluated in 496 cases.

As described above, the minimum information for a diagnostic evaluation to be considered "complete" was a VRA audiogram with reliable MRLs at 1.0, 2.0, and 4.0 kHz for each study ear. If necessary, MRLs were obtained over several sessions. Table 1 shows that hearing status was determined by VRA for 1,184 ears, or 82.7% of the sample. In addition, complete three-frequency audiograms were obtained for the opposite or nonstudy ear in 52.4% of infants.

For a variety of reasons, the audiologists did not always obtain complete and/or definitive diagnostic data for some infants. Sometimes families did not return for a second or third visit, when the infant had failed to condition or had habituated before completing the test at the initial diagnostic session. On occasion, hearing loss due to

Table 1. Completeness of data set.

	Infants enrolled in study	Infants with diagnostic data		Ears	Hearing status determined by VRA at 1.0, 2.0, and 4.0 kHz		Hearing status determined by VRA and OAE		Inconclusive results (<3 frequencies by VRA or OAE)	
		n	%		*n*	%	*n*	%	*n*	%
	1,524	973	63.8	1,432	1,184	82.7	1,284	88.7	148	11.3
Other ear passed				496		52.4		72.2		27.8

Note. VRA = visual reinforcement audiometry; OAE = otoacoustic emission.

transient middle ear effusion prevented getting an estimate of the infant's true hearing status. Thus, for the remainder of study ears, a way was needed to classify those diagnostic results that were not clear-cut. The categories and the process by which they were created are described in detail in White (2005) and summarized below.

For an additional 100 ears, hearing loss was ruled out based on the presence of OAEs at the frequencies where MRLs were not obtained. The remaining 148 study ears had inconclusive data and were further subdivided into categories of "suspicion" or simply "not sufficient data."

In a few instances (5 ears of 3 infants), diagnostic data were obtained from another clinic in the community. This was particularly the case when infants with PHL, who had failed the A-ABR for the *nonstudy* ear, were found to have hearing loss when they were followed by another facility. In most of these instances, the babies had multiple congenital problems such as Down syndrome and cleft palate and had been referred to specialized clinics for follow-up.

Hearing Status of Infants

Since the purpose of this study was to determine how many, if any, of the ears that had failed OAE but passed A-ABR screening had a PHL by 1 year of age, the investigators separated those ears that did not have PHL from those that did. In the 88.7% of ears with definitive data (see Table 1), this was a relatively simple task. However, for the remainder with inconclusive or incomplete results, the investigators reviewed each case and put them in categories related to the likelihood of hearing loss. Table 2 summarizes the hearing status of infants across the categories. (Similar data are also reported in the companion article by White et al., 2005.)

Because numerous case examples will be presented to explain the categories, the format found in Table 3 will be used to display the test results for ease of reading. The VRA MRLs are assumed to be obtained with earphones unless noted otherwise. "Qual" is a code for the examiner's level of confidence in the MRL results (1 = good, 2 = fair). If reliability was considered poor, then MRL was indicated with a "C" for "could not test. A "D" for "did not test" meant that testing was not attempted for the designated frequency. Based on Nozza, Bluestone, Kardatzke, and Bachman (1994) and Roush, Bryant, Mundy, Zeisel, and Roberts (1995), tympanograms were coded as follows: 1 = normal (physical volume between 0.3 and 1.0 cc, static admittance of 0.2 to 1.0 mmhos, and tympanometric width of <235 daPa), 2 = abnormal (normal volume, with static admittance of <0.2 mmhos, tympanometric width of >235 daPa, essentially "no peak" per tympanometric screening), 3 = questionable (one value was missing, usually tympanometric width), and 4 = could not or did not test. OAEs were coded as follows: 2 = SNR < 3 dB, 4 = SNR 3–6 dB, and 6 = SNR $\geq$ 6 dB.

By way of example, Table 3 shows the test results for Case A, who was enrolled in the study because both ears failed OAE and passed A-ABR. At 8 months of age, she was seen for her diagnostic evaluation. For both right and left ears, VRA MRLs were obtained at the levels indicated, suggesting a mild high-frequency hearing loss. The examiner was confident that these MRLs were reliable. Tympanograms were normal. TEOAEs were present at >6 dB SNR from 1.0–3.0 kHz but absent (or < 3 dB) at 4.0 kHz. Similar results were obtained at a second visit

Table 2. Hearing status of infants.

Infants with diagnosis data	Total ears		Not PHL ears		Probable not PHL ears		Ears with PHL				Ears with increased suspicion of PHL				Ears with not sufficient data	
							SN		PC		High		Some			
	n	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%
973	1,432	100	1,140	79.6	100	7	25	1.7	5	0.3	19	1.3	28	2.0	115	8

Note. PHL = permanent hearing loss; SN = sensorineural; PC = permanent conductive.

Table 3. Case example of test results (Case A).

Visit	Age at visit (months)	Right ear MRL 0.5	MRL 1.0	MRL 2.0	MRL 4.0	Qual[a]	Tymp result[b]	OAE[c] 1.0	OAE[c] 2.0	OAE[c] 4.0	Left ear MRL 0.5	MRL 1.0	MRL 2.0	MRL 4.0	Qual[a]	Tymp result[b]	OAE[c] 1.0	OAE[c] 2.0	OAE[c] 4.0
VRA 1	8	15	20	25	35	1	1	6	6	2	20	20	15	25	1	1	6	6	2
VRA 2	11	15	20	25	30	1.5	1	6	6	2	20	20	15	30	1.5	1	6	6	2

Note. Frequencies are kilohertz. Minimum response levels (MRLs) in dB HL.

[a]Quality (level of confidence): 1 = good, 2 = fair. [b]Code for tympanogram: 1= normal, 2 = no peak, 3 = questionable due to missing tymp width, 4 = C or D. [c]Code for OAEs: 2 = signal-to-noise ratio (SNR) < 3 dB, 4 = SNR 3–6 dB, 6 = SNR $\geq$ 6 dB.

when she was 11 months old. A 4-kHz unmasked bone conduction MRL was within 5–10 dB of the air conduction threshold. This infant's hearing status was categorized as sensorineural PHL.

Not PHL. As seen in Table 2, reliable three-frequency (1.0, 2.0, and 4.0 kHz) VRA audiograms ruled out PHL for 1,140 (79.6%) of the study ears. By study definition, ruling out PHL means that MRLs of 15 or 20 dB HL must have been obtained, with good confidence, at 1.0, 2.0, and 4.0 kHz. In fact, most of these results included additional information that could be used as a cross-check of the "normal" results. For example, Case B in Table 4 included MRLs at 500 Hz (also speech awareness thresholds of 15 dB HL that are not shown), tympanometric measures, and OAEs, all indicating normal hearing status. Case C is an example of an infant who was assumed to have a transient conductive hearing loss at the first VRA visit. VRA testing was incomplete, the two MRLs that were obtained were elevated, tympanometry was abnormal, and OAEs were absent. At the second visit, MRLs of 15 and 20 dB HL indicated normal hearing, which was supported by normal tympanometry and OAEs. Case D is an example of an infant who refused earphones on the first visit and was tested successfully on the second visit. Case E is an example of an infant who provided the minimum data to rule out hearing loss.

Probable not PHL. When MRLs were not obtained for the three frequencies but OAEs were present at the frequencies where MRLs were not obtained, it was assumed that there was probably *not* a hearing loss when the OAE SNR was $\geq$3 dB at 1.0 kHz and $\geq$6 dB at 2.0 and 4.0 kHz. In this way, PHL was ruled out in another 100 ears by using a combination of MRLs and OAEs at 1.0, 2.0, and 4.0 kHz. Table 5 shows examples of this category. Case F had OAEs at 1.0 and 4.0 kHz to fill in for the missing MRLs at those frequencies. For Case G, the sound field MRLs were insufficient, but the presence of OAEs at 1.0–4.0 kHz ruled out hearing loss for each ear. Results obtained over two visits provided enough data to rule out hearing loss for Case H. Hearing status for Case I was based on OAEs alone. It could be argued that these criteria were somewhat lenient for ruling out hearing loss. This was intentional. Because the goal of the study was to determine the number of infants with PHL who failed OAE but passed A-ABR at the time of newborn hearing screening, the coinvestigators decided it was a more serious error to designate an infant as having PHL (thus the more rigorous criteria for the PHL group) than to "rule out hearing loss" (thus the more lenient criteria for this "probable not PHL" group).

Inconclusive results. Because of incomplete data, the process of determining the remainder of the ears was not as

Table 4. Case examples: not PHL.

Case	Visit	Age at visit (months)	Right ear MRL 0.5	MRL 1.0	MRL 2.0	MRL 4.0	Qual[a]	Tymp result[b]	OAE[c] 1.0	OAE[c] 2.0	OAE[c] 4.0	Left ear MRL 0.5	MRL 1.0	MRL 2.0	MRL 4.0	Qual[a]	Tymp result[b]	OAE[c] 1.0	OAE[c] 2.0	OAE[c] 4.0
B	VRA 1	8	15	15	15	15	1	1	6	6	6	20	15	15	15	1	1	4	6	6
C	VRA 1	9	C	C	C	C		2	2	2	2	40	C	30	C	1	2	2	2	2
	VRA 2	11	15	15	15	20	1	1	6	6	6	20	15	15	15	1	1	6	6	6
D	VRA 1	8	sf 15	15	15	C	1	3	6	6	6	sf	sf	sf	sf		3	C	C	C
	VRA 2	9	15	15	15	15	1	1	6	6	6						1	6	6	6
E	VRA 1	8	D	20	20	20	1	4	C	C	C									

Note. Frequencies are kilohertz. MRLs in dB HL. D = did not test; C = could not test; sf = sound field (MRLs presented under one ear only).

[a]Quality (level of confidence): 1 = good, 2 = fair. [b]Code for tympanogram: 1= normal, 2 = no peak, 3 = questionable due to missing tymp width, 4 = C or D. [c]Code for OAEs: 2 = SNR < 3 dB, 4 = SNR 3–6 dB, 6 = SNR $\geq$ 6 dB.

Table 5. Case examples: probable not PHL.

			Right ear									Left ear								
			MRL						OAE[c]			MRL						OAE[c]		
Case	Visit	Age at visit (months)	0.5	1.0	2.0	4.0	Qual[a]	Tymp result[b]	1.0	2.0	4.0	0.5	1.0	2.0	4.0	Qual[a]	Tymp result[b]	1.0	2.0	4.0
F	VRA 1	8	15	C	15	D	1	2	4	6	6	15	C	15	20	1	1	6	6	6
G	VRA 1	9	sf	sf	sf	sf		4	6	6	6	sf 20	15	20	20	1	4	6	6	6
H	VRA 1	10	20	C	20	C	1	2	2	2	2									
	VRA 2	11	D	C	D	D		1	6	6	6									
I	VRA 1	8	D	D	C	D		3	6	6	6	D	D	D	D		3	6	6	6

Note. Frequencies are kilohertz. MRLs in dB HL.

[a]Quality (level of confidence): 1 = good, 2 = fair. [b]Code for tympanogram: 1= normal, 2 = no peak, 3 = questionable due to missing tymp width, 4 = C or D. [c]Code for OAEs: 2 = SNR < 3 dB, 4 = SNR 3–6 dB, 6 = SNR $\geq$ 6 dB.

easy as one might think. Ultimately, the investigators reviewed each case that was not clear-cut and placed them into categories of "increased suspicion" or "not sufficient data." From this discussion, the definitions of hearing loss shown (Table 2; White et al., 2005) were developed.

The 115 ears in the "not sufficient data" category were of two primary types. Seventy-five percent were ears with no usable data, that is, MRLs and OAEs were "could not test" or "did not test." See Cases J and K in Table 6 for examples of this category. The remaining 25% were ears for which data appeared normal (MRLs of 15–20 dB HL or present OAEs) but data for one of the three frequencies between 1.0 and 4.0 kHz were missing. In other words, hearing loss at all three target frequencies could not be ruled out, but there was no direct evidence of hearing loss. Because the families did not return as requested, no final determination about hearing status could be made. Examples of these are Cases L and M in Table 6.

There were another 47 ears for which the data were incomplete but the MRLs that were obtained were elevated (>20 dB) and/or the OAEs were absent (<3 dB). This category typically contained MRLs of questionable reliability or abnormal tympanometry. For example, in Table 7, Case N has elevated MRLs but the quality is judged as "fair." Case O had absent OAEs but no data on hearing per se. Over two test sessions, Case P had contradictory findings of elevated MRLs at 4.0 kHz in the right ear and 2.0 kHz in the left, but OAEs >6 dB at each of those frequencies. Case Q had elevated MRLs, abnormal tympanograms, and absent OAEs. The unmasked bone conduction MRLs supported the likelihood of a transient conductive hearing loss.

PHL. Of the ears that failed OAE but passed A-ABR, 30 ears (of 21 infants) were found to have PHL. Their audiometric results are displayed in Table 8. For each ear, MRLs (a composite of the best MRLs obtained across sessions) are shown for 0.5, 1.0, 2.0, and 4.0 kHz. The pure-tone average for 1.0, 2.0, and 4.0 kHz was used to classify the degree of loss, which is given in the adjacent column along with the type of loss. The next section of the table gives details about how the results were obtained. Details about the opposite ear are given where only one ear qualified for the study by failing the OAE and passing A-ABR.

With respect to severity of hearing loss, 23 ears (77%) fell in the mild range, 5 (17%) were moderate, 1 had a severe loss, and 1 had a profound loss. In the mild category, pure-tone averages ranged from 22 to 40 dB HL. Scrutiny of the MRL results at each frequency reveal how some of these ears might have passed an A-ABR screening with

Table 6. Case examples: not sufficient data.

			Right ear									Left ear								
			MRL						OAE[c]			MRL						OAE[c]		
Case	Visit	Age at visit (months)	0.5	1.0	2.0	4.0	Qual[a]	Tymp result[b]	1.0	2.0	4.0	0.5	1.0	2.0	4.0	Qual[a]	Tymp result[b]	1.0	2.0	4.0
J	VRA 1	8	D	C	C	D		1	D	D	D									
K	VRA 1	9	D	D	D	D		3	2	2	2	D	D	D	D		3	C	C	C
L	VRA 1	10	C	C	20	C	1	1	D	6	6	C	C	C	C		1	D	6	6
	VRA 2	10	sf	sf	sf	sf		4	D	6	6	sf 15	15	C	C	1	4	D	6	6
M	VRA 1	8	20	C	15	20	1	4	C	C	C									

Note. Frequencies are kilohertz. MRLs in dB HL.

[a]Quality (level of confidence): 1 = good, 2 = fair. [b]Code for tympanogram: 1= normal, 2 = no peak, 3 = questionable due to missing tymp width, 4 = C or D. [c]Code for OAEs: 2 = SNR < 3 dB, 4 = SNR 3–6 dB, 6 = SNR $\geq$ 6 dB.

Table 7. Case examples: increased suspicion of PHL.

			Right ear									Left ear								
			MRL						OAE[c]			MRL						OAE[c]		
Case	Visit	Age at visit (months)	0.5	1.0	2.0	4.0	Qual[a]	Tymp result[b]	1.0	2.0	4.0	0.5	1.0	2.0	4.0	Qual[a]	Tymp result[b]	1.0	2.0	4.0
N	VRA 1	8	C	C	45	30	2	1	C	C	C									
O	VRA 1	9	D	D	D	D		3	2	2	2	D	D	D	D		3	C	C	C
P	VRA 1	10	sf	sf	sf	sf		1	6	6	6	sf 45	50	D	D	2	1	6	6	6
	VRA 2	10	20	20	15	30	1	1	6	6	6	D	20	25	20	1	1	6	6	6
Q	VRA 1	8	50	50	C	45	1	2	2	2	2	40	D	35	D	1	2	2	2	2
	VRA 2	11	35	D	25	25	1	2	2	2	2	D	40	D	25	1	2	2	2	2
			ub 15	20	15	20	1													

Note. Frequencies are kilohertz. MRLs in dB HL. ub = unmasked bone.

[a]Quality (level of confidence): 1 = good, 2 = fair. [b]Code for tympanogram: 1= normal, 2 = no peak, 3 = questionable due to missing tymp width, 4 = C or D. [c]Code for OAEs: 2 = SNR < 3 dB, 4 = SNR 3–6 dB, 6 = SNR $\geq$ 6 dB.

35-dB nHL click stimuli (as was the case in each of the participating hospitals). One such example is Number 14 in Table 8, who served as the Case A example in Table 3.

Twenty-five of the 30 ears were sensorineural hearing loss, based on elevated MRLs (by either VRA or ABR, and occasional bone conduction MRLs), absence of OAEs, and normal tympanometric measures. Five of the 30 ears (3 infants—2 bilateral and 1 unilateral) appeared to be permanent conductive loss, based on otologic report (i.e., absence of evidence of middle ear effusion, often after myringotomy and tubes) and normal bone conduction responses in the presence of elevated MRL (or elevated tone-burst ABR thresholds). All 3 of the children with permanent conductive hearing loss were from Hawai'i and of Asian or Pacific Islander descent.

Of the 21 infants, 12 had bilateral hearing loss, and 9 had unilateral hearing loss. For 9 of the 12 with bilateral losses, both ears were study ears, that is, both ears had referred on OAEs and passed A-ABR for newborn screening. Three of the 12 infants with bilateral hearing loss had one study ear, but the other ear had failed both OAE and A-ABR for newborn screening (Numbers 3, 4, and 17). Number 17 (also shown in Table 9) was seen at 1 month of age for diagnostic testing because the left ear had failed the A-ABR screen. A profound hearing loss was confirmed for the left ear; OAEs were present at 1.0–3.0 kHz but not at 4.0 kHz in the right (study) ear. However by the 8-month VRA visit, OAE results had changed, and at 9 months, OAEs were absent. ABR revealed a profound hearing loss in the right ear as well as the left, which was later corroborated with corresponding VRA results as shown.

Within the PHL group, the final validating audiometric results were ascertained primarily by the study's VRA protocol in 12 infants (17 study ears). The diagnosis of PHL in 5 ears of 4 infants was ultimately based on tone burst ABR rather than on behavioral thresholds determined by VRA. Besides the infant with the profound loss described above, 1 infant was later diagnosed with autism. Another had Down syndrome with multiple medical problems, as evidenced by a 40-day stay in the NICU.

The infants for whom definitive diagnosis was not made with the study protocol fell into four groups. The diagnosis of PHL in 5 ears of 3 infants was based on diagnostic testing (ABR, OAE, immittance, and VRA MRLs) at another facility. The diagnosis of PHL in 3 ears of 2 infants was based on sound field VRA, not on earphone VRA. In one case, the baby's other ear had an earlier diagnosed loss of greater severity, so the sound field results were assumed to reflect the better ear. The other instance was bilateral symmetrical hearing loss, confirmed by symmetrical ABR thresholds that corresponded with sound field MRLs.

The PHL group was compared with the group with normal hearing (not PHL) on various demographic factors. The only ones that were significantly different statistically are shown in Table 10. The PHL group contained more babies from low-income families or who had low birth weight or had spent time in the NICU than the group of infants who did not have PHL.

Success With the VRA Protocol: Differences Across Sites

Considering the sample as a whole, 88.7 % of the infants were successfully tested, and hearing status was determined using the study protocol. Hearing status was based on VRA MRLs for 82.7% of the total sample. Thirty-two percent of the infants required two visits, 7.6% required three visits, and 7 infants (0.7%) came back a fourth or fifth time.

There are several ways in which "success" with the protocol may be defined. One definition might be a high percentage of ears with complete VRA audiograms. Another might be a high percentage of tests completed in one visit. Within individual test sessions, success may be defined as the ability to condition the infant to do the task or accept earphones.

Just as White et al. (2005) reported site differences in recruitment, referral rates, follow-up rates, and number of infants identified with hearing loss, differences were also noted in how successful the sites were in completing the VRA protocol. Table 11 compares some of these measures for each site. Column 3 shows that success in obtaining

Table 8. Results of diagnostic evaluations of the PHL group.

Record count	Site	ID	Ear	MRL at 0.5 kHz	MRL at 1.0 kHz	MRL at 2.0 kHz	MRL at 4.0 kHz	Mean MRL (1.0, 2.0, and 4.0 kHz)	Type of hearing loss	VRA per study protocol	ABR at site	VRA at other facility	ABR at other facility	VRA sessions	Newborn screening results for other ear	Hearing status of "other" ear
1	2	053	R	25	25	30	35	30	Mild SN							
			L	20	25	30	35	30	Mild SN	sf		x	x	1		
2	2	091	R	20	25	35	35	32	Mild SN							
			L	20	35	35	35	35	Mild SN			x				
3	2	130	L	25	30	35	45	35	Mild SN			x			Ref/Ref	Mod PHL
4	2	131	L	30	35	40	45	40	Mild SN	sf	x	x		1	Ref/Ref	Mild PHL
5	2	148	R	25	30	35	35	33	Mild SN	x				2	Ref/P	NH
6	4	005	R	45	40	40	60	42	Mod PC							
			L	35	35	30	35	33	Mild PC	x				3		
7	4	020	L	35	20	25	20	22	Mild SN	x				5	P OAE	NH
8	4	027	R	20	30	20	40	30	Mild PC							
			L	40	20	50	60	37	Mild PC		x			2		
9	4	029	R	20	20	40	45	35	Mild SN		x			2	P OAE	NH
10	4	055	R	25	25	20	30	25	Mild SN	x				2	P OAE	NH
11	4	066	R	30	40	15	40	32	Mild PC	x				2	P OAE	NH
12	4	089	R	C	35	30	35	33	Mild SN		x			2	Ref/P	Inc susp
13	4	122	R	40	40	45	40	42	Mod SN							
			L	40	40	45	40	42	Mod SN		x			2		
14	4	126	R	15	20	25	30	25	Mild SN							
			L	20	20	15	30	22	Mild SN	x				2		
15	4	138	R	20	30	20	30	28	Mild SN		x					
			L	40	40	30	30	33	Mild SN	x				2		
16	5	046	R	55	65	65	65	65	Severe SN	x				1	Ref/Ref	NH
17	5	065	R	91+	D	91+	91+	91+	Profound SN		x			4	Ref/Ref	PHL
18	6	002	R	25	25	20	20	22	Mild SN							
			R	25	25	30	25	27	Mild SN	x				4		
19	6	003	L	45	45	60	55	53	Mod SN	x				4	Ref/P	NH
20	6	072	L	50	40	30	40	37	Mild SN	x	x			4	Ref/P	NH
21	7	258	R	30	45	30	50	42	Mod SN					2		
			L	40	35	40	40	38	Mild SN	x						

Note. ABR = auditory brainstem response; mod = moderate; ref = refer; P = pass; NH = normal hearing; inc susp = increased suspicion.

Table 9. Case examples: PHL.

			Right ear									Left ear								
			MRL						OAE[c]			MRL						OAE[c]		
Case	Visit	Age at visit (months)	0.5	1.0	2.0	4.0	Qual[a]	Tymp result[b]	1.0	2.0	4.0	0.5	1.0	2.0	4.0	Qual[a]	Tymp result[b]	1.0	2.0	4.0
21	VRA 1	14	30	30	30	45	1	1	2	2	2	35	35	35	40	1	1	2	2	2
(7-258)	VRA 2	15	30	45	30	50	1.5	1	6	6	4	40	35	40	40	1	1	2	2	2
												ub 30	30	40	35	1				
17	Visit 1	1	C	C	C	C		3	6	6	2									
(5-065)	VRA 1	8	D	D	D	D		1	2	6	6									
	ABR		NR		NR	NR		1	2	2	2									
			ub 65	NR	NR	NR														
15	VRA 1	8	ub	D	25	35	1	1	2	2	2	ub	ub	ub	ub	1	1	2	2	2
(4-138)	VRA 2	9	D	45	30	25	1	1	2	2	6	D	30	25	35	1	1	2	2	2
	ABR	17	20	30	20	30		1	2	2	2	40	40	30	30		1	C	C	C
4	VRA 1	9	sf 35	30	45	40	2	1	2	2	2	sf	sf	sf	sf		1	2	2	2
(2-053)	ABR		55		60							55		60						
	VRA	later	25	25	30	35		1	2	2	2	20	25	30	35		1	2	2	2

Note. Frequencies are kilohertz. MRLs in dB HL. NR = no response.

[a]Quality (level of confidence): 1 = good, 2 = fair. [b]Code for tympanogram: 1= normal, 2 = no peak, 3 = questionable due to missing tymp width, 4 = C or D. [c]Code for OAEs: 2 = SNR < 3 dB, 4 = SNR 3–6 dB, 6 = SNR ≥ 6 dB.

complete audiograms with VRA ranged from 91.8% at Site 2 to 54.1% at Site 5. (Differences were statistically significant by one-way analysis of variance, as shown in the bottom row of the table.) Site 2, which was most successful in recruiting and in getting families to return for diagnostic evaluation, was also most successful in obtaining complete three-frequency audiograms for each of the study ears. Adding OAE data to the MRL data (Column 4) allowed all sites to raise the percentage in which hearing status was determined, for example, from 54.1% to 68.5% at Site 5 and to 99.2% at Site 2. Site 1 required the lowest percentage of repeat visits to complete the protocol at 14.8%, whereas Site 4 (which had the highest number of infants identified with hearing loss) required 56.9% of infants to return for at least a second visit. Site 1 was most successful in conditioning with only 4.9% failing to condition on the first visit. Sites 1 and 6 had no trouble getting infants to wear insert earphones, while other sites had more difficulty.

Sites were compared on several factors that might contribute to more successful implementation of VRA. For example, although the VRA protocol was quite structured, alternatives were allowed to accommodate individual preferences of either examiner or infant. For example, examiners were allowed choices in how to begin a test session. Although the protocol required individual ear data for the final analysis, examiners could choose to begin the conditioning process with the sound field presentation of stimuli or they could begin with insert earphones. Likewise, responses to pure tones were required for analysis, but examiners could choose to begin conditioning with speech if they thought that would facilitate conditioning. Table 12 shows that there was considerable difference in how sites chose to begin conditioning. Some sites (1 and 7) always

Table 10. Characteristics of infants with normal hearing and PHL.

Characteristic	% normal hearing in both ears (n = 851)	% PHL in one or both ears (n = 21)	Statistical significance of difference between groups
Family income			$F(1, 870) = 4.74$, $p = .030$
≤$19,999	17.1	23.8	
$20,000–$29,999	11.8	14.3	
$30,000–$49,999	28.2	33.3	
>$50,000	48.9	28.6	
Well baby nursery	87.5	71.4	$F(1, 870) = 4.77$, $p = .029$
Neonatal intensive care unit[a]	12.5	28.6	
Birth weight			$F(1, 870) = 9.39$, $p = .002$
≥2,500 g	87.0	76.2	
≥1,500–2,500 g	7.2	4.8	
1,000–1,500 g	2.6	9.5	
<1,000 g	3.3	9.5	

[a]Two sites did not have NICU participants.

Table 11. Indicators of successful VRA across test sites.

Site	% return for diagnosis evaluation	No. of infants/ears	% complete per VRA	% complete per MRL and OAE	% repeat visits	% failed to condition on Visit 1	% refused earphones on Visit 1
1	42.4	81/148	88.5	92.6	14.8	4.9	0
2	80.8	299/478	91.8	99.2	23.7	10.7	17.1
3	50.0	42/59	67.8	78.0	38.1	26.2	9.5
4	74.1	109/165	58.8	73.3	56.9	10.1	8.3
5	50.6	86/111	54.1	68.5	39.5	19.1	15.6
6	69.2	184/241	85.5	88.8	29.3	13.6	0.5
7	58.1	172/230	85.7	87.8	38.4	11.0	5.8
Statistical significance	$F(6, 972) = 49.13$, $p < .001$	$F(6, 972) = 20.46$, $p < .001$	$F(6, 972) = 16.91$, $p < .001$	$F(6, 972) = 10.15$, $p < .001$	$F(6, 972) = 4$, $p < .001$	$F(6, 972) = 2.41$, $p = .026$	$F(6, 972) = 9.40$, $p < .001$

began with earphones; others (2 and 6) preferred to begin conditioning in the sound field. Most sites began testing with pure tones, but Sites 2 and 5 favored beginning with speech. It appears that how one chooses to begin conditioning has little to do with ultimate success in completing VRA. Sites 1 and 2, with the highest completion, used opposite starting strategies.

Across sites, the mean age at the first VRA session ranged from 8 to 10.9 months. The sites were remarkably alike in the average number of reinforced stimulus trials per session, suggesting that infants everywhere remained on task for a similar number of trials before either completing the test or habituating.

At some sites, the tests for this study were done by the same few examiners; at other sites, tests were done by many different individuals. When the number of testers is compared with the number of infants tested, it becomes apparent that the sites with the fewest examiners per infant were the ones with greater success in implementing the protocol. For example, at Site 1, all tests were done by two audiologists, each sharing the responsibilities at the audiometer or in the room with the child. At Site 2, one audiologist served as Examiner 1 when testing 299 infants. Because there were so many babies at that site, the 12 different test assistants had ample opportunity to gain experience with VRA testing. Site 4, on the other hand, used 5–6 different testers to test 109 babies, thus tester time per baby was limited. Site 5 used nine examiners to test its 86 study babies. Sites 4 and 5 had the lowest percentages of complete tests. It is apparent that practice and experience play an important role in success with VRA.

Discussion

Although the focus of this article is on the diagnostic behavioral testing of the infants enrolled in the study, the prerequisite to diagnostic testing—follow-up of recommendations from screening—is highlighted here. Only 64% of the families enrolled in the study returned for diagnostic evaluation, despite postcard communication in the intervening months, considerable time and effort of study personnel to locate the families and schedule evaluations, and reimbursement for costs of time and travel. The same percentage returned for the NIDCD/INHI study with similar time intervals, staff effort, and inducements (Norton et al., 2000). Some of the possible reasons for *not* returning are outlined in White et al. (2005). The follow-up rates for research studies such as these are not very different from the follow-up rates reported for universal newborn hearing screening programs across the United States. With screening failure rates as low as 1%–2%, there is a high probability that an infant who is referred for diagnostic testing has a hearing loss. The role of the primary care physician is crucial in helping families understand the importance of following up on the referral from newborn hearing screening.

The generally accepted goal of universal newborn hearing screening is to diagnose hearing loss before 3 months of age so that intervention is begun before 6 months. A protocol based on VRA, like the one used in this study, will not be appropriate for such early diagnosis. However, such a protocol *is* appropriate for infants with risk factors for later onset hearing loss and for the continued detailed

Table 12. Differences in implementation of VRA protocol across sites.

Site	No. of infants/ears	% sessions begun with earphones	% sessions begun with pure tones	Mean age in months at VRA 1	Mean reinforced stimulus trials per session	No. of audiologists serving as Examiner 1	Total number of examiners
1	81/148	100	100	10.6	28.3	2	2
2	299/478	29	0	8.0	39.9	1	13
3	42/59	80	98	10.1	29.3	3	3
4	109/165	90	81	10.9	29.5	5	11
5	86/111	47	7	9.6	33.2	1	9
6	184/241	7	99	8.4	33.9	3	5
7	172/230	99	99	9.8	32.8	2	11

audiometric testing of infants with congenital hearing loss. From this study, we found that for the majority of 8–12-month-old infants, hearing sensitivity can be reliably assessed using VRA. Two thirds can be tested behaviorally, with insert earphones, in a reasonable length of time (20–30 min) in one test session. Nevertheless, our results also suggest that behavioral testing will often take more than one session to complete. In our sample, nearly 1 in 3 infants had to return for continued assessment.

Conversely, some infants were difficult to test, requiring multiple visits to definitively establish hearing status. For infants with hearing loss, more visits may result in delays in diagnosis, fitting of amplification, and enrollment in early intervention programs. Some infants with severe hearing loss and/or other disabilities did not respond to behavioral audiometry but required ABR for the definitive diagnosis of type and severity of hearing loss. Supplementary physiological measures such as OAEs and tympanometry are important complements to test validity. Audiologists may be more comfortable making a final diagnosis when they are able to corroborate the behavioral findings with results that do not require a voluntary response from the infant, but it is important that such corroborating procedures not unduly delay the determination of hearing status.

The factors that may influence the need for multiple test sessions may include audiologists' unwillingness to "trust" their results. Even experienced pediatric audiologists are not equally successful in implementing a VRA protocol in their clinical situations. The findings of this study suggest that experience helps; that is, successful use of a structured VRA protocol improves as examiners test more babies, more often. The findings also suggest that details of conditioning, such as type of beginning stimulus and transducer, do not influence the ultimate success of the VRA test session. Understanding the concept of having an infant under stimulus control may help audiologists accept that when conducted in a controlled manner, VRA measures are not subjective. A possible disadvantage of examiners' having greater confidence in "normal" results is that examiners may miss diagnosing hearing loss, particularly mild hearing loss. A protocol that incorporates the use of probe and control trials as measures of stimulus control should help to alleviate possible problems. Likewise, a cross-check with physiological measures is equally important for the determination of normal hearing sensitivity.

Another lesson exemplified in this study is that hearing status of infants may change over time. Unilateral and OAE screening failures may warrant closer monitoring. When a newborn fails hearing screening in one ear, it behooves the audiologist to include the other passing ear in rescreening and diagnostic testing. Several infants in our study were found to have mild losses in the ear that was assumed to be the good ear. At least one of our study infants also appeared to have a rapidly progressing loss in the better ear that passed newborn screening. This study also lends support to other studies which show that hearing levels can be elevated substantially in the presence of abnormal tympanograms suggestive of middle ear effusion and that hearing sensitivity can be very different from one day to another (Gravel & Wallace, 2000; Sabo et al., 2003). Thus, it is imperative that parents and professionals realize that passing the newborn hearing screening test does not mean that hearing loss may not occur at a later time.

Acknowledgments

This study was made possible through a Cooperative Agreement between the Centers for Disease Control and Prevention (CDC) and the Association of Teachers of Preventive Medicine (ATPM), Award U50/CCU300860 TS-523. The contents of the study are the sole responsibility of the authors and do not necessarily reflect the official views of the CDC or the ATPM. An abbreviated version of the four articles in this series was previously published in the September 2005 issue of *Pediatrics* (Johnson, White, Widen, Gravel, James, et al., 2005). The contributed resources of the National Center for Hearing Assessment and Management were indispensable in the successful completion of the study. Justus Randolph, Danhui Zhang, and Maria Hovak were critical assets in the management and analysis of the data. Three audiology graduate students at the University of Kansas contributed to preparation of this particular article: Stacey Baldwin, Rosa Gutierrez Moore, and Angelique Torres. The authors express their heartfelt appreciation to the pediatric audiologists who performed the diagnostic assessments. Finally, we thank the 973 families who returned for the diagnostic evaluations.

References

American National Standards Institute. (1996). *Specification for audiometers (ANSI S3.6-1996).* New York: Author.

Day, J., Bamford, J., Parry, G., Shepherd, M., & Quigley, A. (2000). Evidence on the efficacy of insert earphone and sound field VRA with young infants. *British Journal of Audiology, 34,* 329–334.

Gravel, J., Berg, A., Bradley, M., Cacace, A., Campbell, D., Dalzell, L., et al. (2000). New York State Universal Newborn Hearing Screening Demonstration Project: Effects of screening protocol on inpatient outcome measures. *Ear and Hearing, 21,* 131–140.

Gravel, J. S., & Wallace, I. F. (2000). Effects of otitis media with effusion on hearing in the first 3 years of life. *Journal of Speech, Language, and Hearing Research, 43,* 631–644.

Gravel, J. S., White, K.R., Johnson, J. L., Widen, J. E., Vohr, B. R., James, M., et al. (2005). A multisite study to examine the efficacy of the otoacoustic emission/automated auditory brainstem response protocol: Recommendations for policy, practice, and research. *American Journal of Audiology, 14,* S217–S228.

Johnson, J. L., White, K. R., Widen, J. E., Gravel, J. S., James, M., Kennalley, T., et al. (2005). A multicenter evaluation of how many infants with permanent hearing loss pass a two-stage otoacoustic emissions/automated auditory brainstem response newborn hearing screening protocol. *Pediatrics, 116,* 663–672.

Johnson, J. L., White, K. R., Widen, J. E., Gravel, J. S., Vohr, B. R., James, M., et al. (2005). A multisite study to examine the efficacy of the otoacoustic emission/automated auditory brainstem response newborn hearing screening protocol: Introduction and overview of the study. *American Journal of Audiology, 14,* S178–S185.

Joint Committee on Infant Hearing. (1994). Joint Committee on Infant Hearing 1994 position statement. *Pediatrics, 94,* 152–156.

Lutman, M. E., Davis, A. C., Fortnum, H. M., & Wood, S. (1997). Field sensitivity of targeted neonatal hearing screening by transient-evoked otoacoustic emissions. *Ear and Hearing, 18,* 265–276.

Mason, S., Davis, A., Wood, S., & Farnsworth, A. (1998). Field sensitivity of targeted neonatal hearing screening using the Nottingham ABR screener. *Ear and Hearing, 19,* 91–102.

Norton, S., Gorga, M., Widen, J., Folsom, R., Sininger, Y., Cone-Wesson, B., et al. (2000). Identification of neonatal hearing impairment: Evaluation of transient evoked otoacoustic emission, distortion product otoacoustic emission, and auditory brain stem response test performance. *Ear and Hearing, 21,* 508–528.

Nozza, R. J., Bluestone, C. D., Kardatzke, D., & Bachman, R. (1994). Identification of middle ear effusion by aural acoustic admittance and otoscopy. *Ear and Hearing, 15,* 310–323.

Roush, J., Bryant, K., Mundy, M., Zeisel, S., & Roberts, J. (1995). Developmental changes in static admittance and tympanometric width in infants and toddlers. *Journal of the American Academy of Audiology, 6,* 334–338.

Sabo, D. L., Paradise, J. L., Kurs-Lasky, M., & Smith, C. G. (2003). Hearing levels in infants and young children in relation to testing technique, age group, and the presence or absence of middle-ear effusion. *Ear and Hearing, 24,* 38–47.

Shannon, D. A., Felix, J. K., Krumholz, A., Goldstein, P. J., & Harris, K. C. (1984). Hearing screening of high risk newborns with brainstem auditory evoked potentials: A follow-up study. *Pediatrics, 73,* 22–26.

Swigonski, N., Shallop, J., Bull, M. J., & Lemons, J. A. (1987). Hearing screening of high risk newborns. *Ear and Hearing, 8,* 26–30.

Talbott, C. B. (1987). A longitudinal study comparing responses of hearing-impaired infants to pure tones using visual reinforcement and play audiometry. *Ear and Hearing, 8,* 175–179.

van Straaten, H. L. M., Groote, M. E., & Oudesluys-Murphy, A. M. (1996). Evaluation of an automated auditory brainstem response infant hearing screening method in at risk neonates. *European Journal of Pediatrics, 155,* 702–705.

Vohr, B. R., Oh, W., Steward, E. J., Bentkover, J. D., Gabbard, S., Lemons, J., et al. (2001). Comparison of costs and referral rates of 3 universal newborn hearing screening protocols. *Pediatrics, 139,* 238–244.

Watson, D. R., McClelland, R. J., & Adams, D. A. (1996). Auditory brainstem response screening for hearing loss in high risk neonates. *International Journal of Pediatric Otorhinogology, 36,* 147–183.

White, K. R., Vohr, B. R., Meyer, S., Widen, J. E., Johnson, J. L., Gravel, J. S., et al. (2005). A multisite study to examine the efficacy of the otoacoustic emission/automated auditory brainstem response newborn hearing screening protocol: Research design and results of the study. *American Journal of Audiology, 14,* S186–S199.

Widen, J., Folsom, R., Cone-Wesson, B., Carty, L., Dunnell, J., Koebsell, K., et al. (2000). Hearing status at 8 to 12 months corrected age using a visual reinforcement audiometry protocol. *Ear and Hearing, 21,* 471–487.

Wilson, W. R., & Thompson, G. (1984). Behavioral audiometry. In J. Jerger (Ed.), *Pediatric audiology*. San Diego, CA: College Hill Press.

Received May 7, 2005
Accepted November 8, 2005
DOI: 10.1044/1059-0889(2005/022)

Contact author: Judith Widen, Department of Hearing and Speech, MS 3039, University of Kansas Medical Center, 3901 Rainbow Boulevard, Kansas City, KS 66160. E-mail: jwiden@kumc.edu

Judith S. Gravel is now at The Children's Hospital of Philadelphia.

Appendix A

Association of Teachers of Preventive Medicine (ATPM) VRA Protocol Summary

Pulsed, warbled tones of 1–2 seconds duration.
Vary ISI, longer ISI initially if random head turns are frequent.

Note beginning time

BEGIN with 2.0 kHz warble tone with earphones
(Examiners may choose to begin with speech and/or sound field).
35 dB HL … if baby turns naturally
2 correct consecutive responses on its own, → go to *TEST PROTOCOL*

35 dB … no head turn
55 dB … if turn, reinforce
2 consecutive responses, go to *TEST PROTOCOL*
…if no turn, go to *CONDITIONING*

CONDITIONING TRIALS
55 dB paired with reinforcer, 2 times
then Probe (reinforce only if child turns)
2 consecutive head turns on its own before going to *TEST PROTOCOL*

No turn on probe? Go to 75 dB HL, pair with reinforcer
2 times
then Probe (reinforce if child turns; get 2 responses on own as above)

Hearing problem or Conditioning problem?
Increase level? Change stimulus? Change ear? Try sound field?

TEST PROTOCOL … after 2 consecutive head turns on own
Down 20, up 10 dB for first reversal; then down 10, up 5 thereafter.
Insert control trials according to Worksheet schedule (before or after stimulus trial)
15 dB HL (or lowest level)
2 times (or MRL = lowest level with 2 responses out of 3)

2nd frequency = 500 Hz
Begin at level of previous response, i.e., 15 dB
2 times, etc.
3rd frequency = 4.0 kHz
4th = 1.0 kHz

Speech may be inserted at any place in the protocol at examiner's discretion.
SAT should be obtained using the same protocol as tones
SAT = lowest level at which 2 of 3 responses are obtained.

Second ear
35 … if head turn (either side, reinforce on side of turn)
(1 time is enough), drop to 15 and proceed as before
… if no head turn, proceed as with first ear, i.e., to 55, then pair

Deviations from this order may arise if child begins to habituate
- change stimuli (speech or WT) at 15 dB HL or level of last response
- re-conditioning trials at level child responded to previously, 1 paired trial, then probe; if child turns, down 20, up 10 again; if no turn, proceed as with initial conditioning trials.

Note End Time

Appendix B

ATPM VRA Protocol Worksheet

Name:______________________________ Date:____________________

Go across with each increase in level. Go down when level is decreased or repeated.
Add control during sequence when "C" occurs on that line.

"+" = head turn
"o" = no head turn
"P+" = Paired conditioning trial with head turn
"Po" = Paired conditioning trial with no head turn

Test Sequence: 2.0, 0.5, 4.0, 1.0 kHz, Speech > anywhere in the sequence.

Begin Time:____________ End Time:____________ No. of breaks:____________

			RESPONSE LEVEL (dB HL)												
EAR	STIM TYPE	CON	15	20	25	30	35	45	55	65	75	85	95		COMMENTS
		C													
		C													
		C													
		C													
		C													
		C													
		C													
		C													
		C													
		C													
		C													
		C													
		C													

Appendix C (p. 1 of 2)

ATPM Behavioral Evaluation No. 1

Name_____(optional) Enrollment # __ __ __ __ __

1. Date of test __ __ /__ __/__ __ (mo dy yr) 2. Tester initials __ __ __ Assistant initials __ __ __

3. Neonatal test results: Right Ear ❑ Pass OAE ❑ Refer OAE/Pass ABR ❑ Refer OAE/Refer ABR
Left Ear ❑ Pass OAE ❑ Refer OAE/Pass ABR ❑ Refer OAE/Refer ABR

History Information Since Neonatal Screening

4. ❑Y ❑N ❑U Parental concern regarding hearing status

5. ❑Y ❑N ❑U Infant has known hearing loss If yes, ❑ Right Ear ❑Left Ear
If yes, infant has hearing aid(s) ❑Y ❑N ❑U

6. ❑Y ❑N ❑U Family history of childhood hearing loss
(Include infant's parents, siblings, grandparents, aunts, uncles, first cousins; must be blood relation to infant)
If yes, specify type of hearing loss and relationship to patient
_____

7. ❑Y ❑N ❑U Congested today

8. ❑Y ❑N ❑U Ear infections diagnosed by a physician
If yes, number of infections to date __ __ or ❑U
infection within last month ❑Y ❑N ❑U
ear tubes at time of this test ❑Y ❑N ❑U

9. ❑Y ❑N ❑U Illness/condition/treatment since screening possibly affecting hearing or neurological status
(Including herpes, meningitis, pertussis, mumps, measles, immunization reaction, sepsis, tuberculosis, hypoxia, seizures, hydrocephalus, head trauma, chemotherapy, radiation therapy, kidney failure, cystic fibrosis, mechanical ventilation, ECMO, cardiac surgery, neurological illness, neurological surgery,or newly diagnosed syndrome)
If yes, specify _____

Minimum Response Levels

		(For each level, indicate response or ✓ one box)				Confidence in MRL value
10. Right Ear	500	__ __ __ dB HL	❑ No response	❑ CNT	❑ DNT	❑ Good or ❑ Fair
	1000	__ __ __ dB HL	❑ No response	❑ CNT	❑ DNT	❑ Good or ❑ Fair
	2000	__ __ __ dB HL	❑ No response	❑ CNT	❑ DNT	❑ Good or ❑ Fair
	4000	__ __ __ dB HL	❑ No response	❑ CNT	❑ DNT	❑ Good or ❑ Fair
	SAT	__ __ __ dB HL	❑ No response	❑ CNT	❑DNT	❑ Good or ❑ Fair
11. Left Ear	500	__ __ __ dB HL	❑ No response	❑ CNT	❑ DNT	❑ Good or ❑ Fair
	1000	__ __ __ dB HL	❑ No response	❑ CNT	❑ DNT	❑ Good or ❑ Fair
	2000	__ __ __ dB HL	❑ No response	❑ CNT	❑ DNT	❑ Good or ❑ Fair
	4000	__ __ __ dB HL	❑ No response	❑ CNT	❑ DNT	❑ Good or ❑ Fair
	SAT	__ __ __ dB HL	❑ No response	❑ CNT	❑ DNT	❑ Good or ❑ Fair
12. Sound field	500	__ __ __ dB HL	❑ No response	❑ CNT	❑ DNT	❑ Good or ❑ Fair
	1000	__ __ __ dB HL	❑ No response	❑ CNT	❑ DNT	❑ Good or ❑ Fair
	2000	__ __ __ dB HL	❑ No response	❑ CNT	❑ DNT	❑ Good or ❑ Fair
	4000	__ __ __ dB HL	❑ No response	❑ CNT	❑ DNT	❑ Good or ❑ Fair
	SAT	__ __ __ dB HL	❑ No response	❑ CNT	❑ DNT	❑ Good or ❑ Fair

Test Specifics

14. Beginning stimulus type (✓ one) ❑ Warble tone ❑ Speech

15. Beginning transducer (✓ one) ❑ Earphone ❑ Speaker

16. # Beginning conditioning trials __ __ (Include P_0 and P^1 before first descent in session; do not include probe trials)

Appendix C (p. 2 of 2)

ATPM Behavioral Evaluation No. 1

17. # Reconditioning trials __ __ (Include P_0 and P^+ trials *after* first descent in session)

18. # Stimulus trials __ __ __ (Include P_0, P^+, + and 0 trials; do not include controls)

If > 0, # of reinforced __ __ __ (Include P^+ and + trials)

19. # Control trials __ __ (Include C^+ and C_0 trials)

If > 0, # correct __ __ (Include C_0 trials only) % of controls correct _____ (Leave blank for data entry)

20. Test time in minutes ___ ___ 21. # Breaks taken __ __

Immittance Results

22. RIGHT ear evaluated ❑Yes ❑ CNT ❑ DNT

If yes, Ear canal volume __ . __ ml

Peak pressure __ __ __ __ daPa ❑ No peak

Admittance __ . __ ml ❑ No peak

Tymp width __ __ __ daPa ❑ No peak ❑ Not available

23. LEFT ear evaluated ❑Yes ❑ CNT ❑DNT

If yes, Ear canal volume __ . __ ml

Peak pressure __ __ __ __ daPa ❑ No peak

Admittance __ . __ ml ❑ No peak

Tymp width __ __ __ daPa ❑ No peak ❑ Not available

OAE Results

24. Type : ❑ TEOAE ❑ DPOAE

25. Right ear

OAE SNR at 1–1.9 kHz	❑ < 3dB	❑ ≥ 3 dB < 6 dB	❑ ≥ 6 dB	❑ CNT	❑ DNT
OAE SNR at 2–2.9 kHz	❑ < 3dB	❑ ≥ 3 dB < 6 dB	❑ ≥ 6 dB	❑ CNT	❑ DNT
OAE SNR at 3–3.9 kHz	❑ < 3dB	❑ ≥ 3 dB < 6 dB	❑ ≥ 6 dB	❑ CNT	❑ DNT
OAE SNR at 4–4.9 kHz	❑ < 3dB	❑ ≥ 3 dB < 6 dB	❑ ≥ 6 dB	❑ CNT	❑ DNT

26. Left ear

OAE SNR at 1–1.9 kHz	❑ < 3dB	❑ ≥ 3 dB < 6 dB	❑ ≥ 6 dB	❑ CNT	❑ DNT
OAE SNR at 2–2.9 kHz	❑ < 3dB	❑ ≥ 3 dB < 6 dB	❑ ≥ 6 dB	❑ CNT	❑ DNT
OAE SNR at 3–3.9 kHz	❑ < 3dB	❑ ≥ 3 dB < 6 dB	❑ ≥ 6 dB	❑ CNT	❑ DNT
OAE SNR at 4–4.9 kHz	❑ < 3dB	❑ ≥ 3 dB < 6 dB	❑ ≥ 6 dB	❑ CNT	❑ DNT

Follow-up

27. ❑Y ❑N Schedule additional behavioral evaluation session for ATPM Study. If yes, indicate reason (✓ all that apply)

❑ Failed to condition ❑ Fussy/scared/tired ❑ Wax occluding canal ❑ Refused earphones

❑ Habituated before completing test ❑ Test reliability not good ❑ Ran out of time

❑ Abnormal tympanogram(s) and MRL(s) > 20 dB HL ❑ MRL(s) > 20 dB HL; bone cond testing needed

28. ❑Y ❑N Medical follow-up recommended.