WHITE PAPER

Automated ABR (AABR):

Considerations for Technology Selection and Meeting the Standard of Care

Authors:
Aaron Thornton, PhD[1]
Judith A. Marlowe, PhD, FAAA, CCC-A[2]

The rapid adoption of universal newborn screening for early detection of hearing loss over the past two decades has dramatically increased the number of infants being tested prior to hospital discharge. From the earliest days, personnel requirements to staff such programs fueled interest in objective measures with 'automated' results in order to allow existing staff or non-technical employees to operate screening equipment reliably and accurately.

Current screening recommendations

In their 2000 Position Statement, the Joint Committee on Infant Hearing (JCIH) provided the first specific guidance on screening technology choices in early hearing detection programs:

> "Programs that use trained and supervised nonprofessional staff must use technologies that provide automated pass-refer criteria. **Before incorporating automated response detection algorithms, however, the screening program must ensure that the algorithms have been validated by rigorous scientific methods and that those results have been reported in peer-reviewed publications**." (p.803) *[Emphasis added]*

The JCIH 2007 Position Statement reaffirmed that automated ABR and otoacoustic emissions (OAE) are the two acceptable physiologic measures for screening, recognized the increased number of available equipment options, and emphasized the responsibility of screening facilities to ensure careful consideration when making their choice:

> "Hospital-based programs should consider screening technology (i.e., OAE or automated ABR testing); **validity of the specific screening device**; screening protocols, including the timing of screening relative to nursery discharge; availability of qualified screening personnel; suitability of the acoustical and electrical environments; follow-up referral criteria; referral pathways for follow-up; information management; and quality control and improvement.
>
> Interpretive criteria for pass/fail outcomes should reflect clear scientific rationale and should be evidence based. Screening technologies that incorporate automated- response detection are necessary to eliminate the need for individual test interpretation, to reduce the effects of screener bias or operator error on test outcome, and to ensure test consistency across infants, test conditions, and screening personnel.
>
> **Manufacturers of hearing-screening devices do not always provide sufficient supporting evidence to validate the specific pass/fail criteria and/or automated algorithms used in their instruments**." (pp. 903-904) *[Emphasis added]*

Introducing automated ABR detection

Auditory Brainstem Response (ABR) is the most accurate technique for assessing auditory function in newborns and has been regarded as the gold standard for diagnostic evaluation during the first six months of life for more than 30 years (Durieux-Smith A et al, 1985; Fria, 1985; Galambos R, Hicks GE, Wilson MJ, 1984; Hall JW et al, 1988; Hyde M et al, 1984; Jacobson JT, Jacobson CA, Spahr RC, 1990; Stein, 1984). By determining the presence or absence of the ABR at various intensity levels, hearing threshold sensitivity may be estimated. In traditional methodology, the EEG is sampled, averaged, and waveforms are recorded for stimuli at decreasing intensities to determine the lowest level that evokes a detectable response. Skilled clinical interpretation is required to correctly identify the presence of the ABR and to distinguish it from physiologic noise, however. The strength of ABR as an objective physiologic measure is compromised by subjective interpretation and the opportunity for human error. For this reason, a new approach to ABR was required to make universal newborn hearing screening feasible.

Throughout the past twenty years, medical devices using automated ABR detection have been developed and have received regulatory clearance for clinical use. The ALGO newborn hearing screener with Natus Medical's proprietary, patented ALGO AABR was the first fully automated ABR technology to be developed with critical design parameters that have carried forward into the current family of ALGO screeners. It uses a binomial template-matching detector that provides exact control over the error of passing a deaf baby. By sampling and processing data only when the baby is resting quietly enough to detect the smallest response that might occur, ALGO also keeps hearing babies from failing the test and from needing unnecessary and expensive follow up.

natus®

The ALGO screener was introduced first in 1985 after clinical research by audiologists determined the feasibility of using automated ABR detection for accurate, low-cost newborn hearing screening. This was accomplished by examining the limits of ABR detectability and reliability; optimizing ABR performance through the creation of an ABR-based hearing screener specifically designed for infants in the nursery environment and automating the technology to ensure reliable "hands-off" operation by available personnel (Herrmann BS, Thornton AR, Joseph JM, 1995). The details of its detection algorithm are described in ALGO Clinical Series No. 3: Inside the Black Box—*How Does The ALGO® Work Anyway?*

The ALGO screener is designed for the single purpose of ABR screening in newborns, and the investigations that contributed to design decisions were intended to identify, optimize, and automate those factors most important in ABR response detection at low stimulus levels. A template derived from the responses of normally hearing infants presented with 35dBnHL click stimuli was incorporated into the statistical response detection method employed.

Numerous independent, peer-reviewed reports describing the performance of ALGO AABR have appeared in the clinical literature since its introduction. These publications range from validation studies with selected cohorts to population-based field reports from state and national hospital based newborn screening programs in the U.S. and abroad documenting the clinically proven sensitivity and specificity of ALGO screeners (See Bibliography).

What do automated ABR screeners detect?

In the simplest definition, the ABR is the neurogenic electrical activity from the auditory nerve and lower portion of the neural pathway to the primary auditory cortex of the brain in response to a brief audible sound. The ABR is typically recorded from electrodes attached to the head and has a duration of 1/30 second or less, depending upon the characteristics of the sound and the parameters of the recording instrument. The ABR stimulated by a regular series of loud click stimuli is often shown as waveform having five positive peaks in the first 1/100 second from the onset of the stimulus, and that characteristic is remarkably consistent from person to person. At slow stimulus rates, the ABR waveform is reasonably interpreted as the waveform that would be observed following each stimulus.

Changes in stimulus intensity, stimulation rate, as well as spectral and temporal properties of the stimuli, electrode location, and processing of the raw data can alter markedly the appearance of the ABR waveform. At high stimulation rates, responses from earlier and later portions of the auditory pathway begin to overlap, even by several times. Under these conditions, the conventionally averaged ABR waveform is no longer a valid estimate of the waveform in response to a single stimulus.

Of greater concern is adaptation of the ABR at high stimulus rates, which effectively sets an upper rate limit beyond which the time saved by increasing rate is offset by decline in the size of the ABR. The ALGO screener uses a slower stimulus rate that permits evaluation along the entire duration of the ABR to a single stimulus, and the multipoint template provides the equivalent of three independent samples per stimulus. This matches or exceeds the effective rate of information acquired by higher-rate alternative methods, which are also more susceptible to periodic artifact at the stimulation rate.

Comparing methods

The ABR to a single stimulus is much smaller than the random physiologic noise, mostly from muscle activity (EMG) picked up by the recording electrodes, so the ABR cannot be directly observed. Detection methods depend upon combining the data from many stimulus presentations since the random EMG contribution cumulates more slowly than that from the synchronous ABR. Most processing methods estimate variability of the EMG from the sampled data and use that to determine whether the combined data is larger than would be predicted by the EMG alone. This can be done by variance analysis in the time or frequency domains using any number of models.

All methods for detecting signals (ABR) in the presence of physiologic noise (EMG) are based on theoretical statistical models that define their expected performance, providing that all theoretical assumptions in the model prove to be valid in real world application. When assumptions cannot be met, the theory is invalidated and the accuracy of the method cannot be defined from its theory.

The detection model used by the ALGO screener is very robust regarding the degree to which important assumptions about physiologic noise are violated, as happens frequently with fussy babies. Because the binary detection model used by the ALGO precisely quantifies the probability of the EMG being mistaken for ABR from a theoretical distribution, the specific characteristics of EMG noise have no effect on the probability of passing a deaf baby, the most serious error for screening (Schwartz and Shaw, 1975).

Other methods that average the amplitude of the electrical activity must assume not only that the EMG has a normal amplitude distribution similar to a bell curve, but that the EMG remains stationary throughout the test. When that isn't true, there is no correction that can be applied for the failure of the model to accurately compute the probability of passing a deaf baby. Under adverse conditions there may be an unknowable increase in the likelihoods of failing a hearing baby, a correctible, but expensive and emotional error and / or passing a deaf baby, the most serious error. Passing a deaf baby, may not be corrected for months or years. Given the relatively low incidence of deafness and delays in learning of a screening identification failure, it may take a very long time for a program to ascertain the accuracy of a screening method from direct experience.

Detection models also differ in their sensitivity to the size of the ABR relative to that of the EMG, known as the Signal to Noise ratio or S/N. Each model has an absolute lower limit to the S/N needed to detect a response, regardless of how many stimuli are presented or how long the test becomes. Because the sizes of the ABR and the EMG fluctuate during the course of running a test, the number of stimulus presentations needed to detect a response is impossible to set at the start of a test. The ALGO detection model provides a direct measurement of S/N, so it can accurately pass a baby as soon as the S/N exceeds a stringent statistical criterion set to ensure against a false detection. This can occur very rapidly when the ABR is large and the EMG is small.

ALGO refers a baby when no response is detected after it has processed enough data to have detected a very small response measured when the EMG was below a criterion level. Thus, the time needed to accurately refer a baby is conditional on how quietly the baby is resting. The size or other characteristics of the EMG noise is often the most important factor in ABR detectability, and all ABR screeners make some attempt to limit ABR processing to those time intervals that the EMG noise is below a level thought to permit an accurate test. But the specific methodology for doing this varies greatly as does the effectiveness. The size of the ABR not only depends upon hearing sensitivity but also on audibility. Background sounds, common in the screening environment, may mask out the audibility of the test stimuli, reducing the size of the ABR during the time that the noise is present. Screener performance will depend upon how well it recognizes and manages these problems.

ABR hearing screeners can be difficult to compare due to the multiple factors that contribute to their operational costs, ease of use and accuracy. While they often approach stimulation and ABR detection in different ways, they are all subject to the same fundamental principles of hearing and electrophysiology of the auditory pathway. For a given stimulation of the ear, there is a limited amount of information encoded in the ABR, and the challenge for any screener is to best use that information to enhance detection and control the probability of error.

Detection models are important, but they do not overshadow the quality of implementation. The ALGO was refined well before its introduction, and that process of improvement has continued, incorporating knowledge gained in 30 years of experience.

Informed choice and quality assurance

Differences in ABR detection methods and implementation mean that all ABR hearing screeners cannot be considered equal and that understanding the specific response detection method employed in a screener must include weighing the theoretical operation under perfect conditions against performance in less than ideal circumstances where the theory may not be valid. The ability to adjust screening parameters and "review" waveforms or "confirm results" may have a significant impact on reliability, consistency and cost in newborn hearing screening programs. Unfortunately, these issues often are not addressed by manufacturers reporting clinical performance. Even more concerning, sensitivity and specificity evidence is rarely (if ever) provided in writing.

The advantage of ALGO AABR is its fixed, fully-automated technology, which cannot be adjusted and which does not require review or confirmation. Early after the ALGO was introduced, there were experienced users who demanded being able to adjust and override parameters of operation. But whenever that was granted, testing errors increased dramatically. ALGO AABR has been clinically proven in diverse settings worldwide for more than three decades with independent, peer-reviewed, clinical publications as evidence of its effectiveness and efficiency in early hearing detection.

Implications for selection of technology

EHDI managers must be able to assess clinical literature as well as product claims made by commercial representatives of a technology in order to confirm that there is adequate evidence supporting the detection algorithm used by the specific device and that the reliability of its performance is well-documented.It is unwise to assume because a bibliography is available or because a report appears authoritative that publications are relevant and answer critical questions necessary to select accurate, reliable technology. Criteria proposed by Silverman (1993) apply equally when considering both published reports and advertised technology claims:

Validity

Does the data presented refer directly to the response detection method used in the current version of device? Do studies or program results offered as evidence represent the actual detection method or do they describe modified screening parameters or adapted response detection methods? Are sensitivity and specificity based on validation against independent measures considered the "gold standard" for the specific test method reported? Clinical publications or product literature providing only refer rates do not indicate the accuracy of the device's response detection method unless follow-up status of passed infants as well as those referred is included.

Reliability

Is the data presented repeatable, or has it only been observed or reported in limited fashion? The number of subjects referenced is critical as a small sample is less likely to be representative. Was data derived in a manner that avoided random error and investigator bias? When same or similar conclusions have been reached by different independent investigators, the more reliable they are. When publications are written predominantly by affiliates of the manufacturing company, the likelihood of bias is significant.

Generality

Were the testing environment and screening personnel representative of the intended use of the screener? Data collected in acoustically controlled clinical settings will not reflect noise and other interference typical of newborn nurseries and mother's room. Likewise, screening conducted by clinical audiologists, ENT physicians, or other practitioners may not be relevant in an institution planning to rely upon technicians or volunteers to operate the screener. Do the subjects of the report represent the intended population? This is most likely to be the case if data and conclusions are based upon a random sample or consecutive cases rather than a selected sample. Results reported as "ears" rather than "infants" may not reflect the real life performance of a technology. A particular concern is whether subjects are newborns age 24-48 hours when screening typically is performed. Or does the publication report results of older infants? In the latter case, attempts to generalize the data may yield only an uncertain estimate of actual results in the pre-discharge infant.

Essential regulatory questions to answer

Evaluating and selecting technology is a process that should identify and minimize risks associated with device misuse or failure for the patient and the institution. As the principal consumer protection agency of the US Federal Government, the US Food and Drug Administration (FDA) has established regulations for the manufacture and sale of medical devices nationwide. Important issues related to safety and effectiveness require knowledge of the specific device with respect to FDA:

1. **FDA status of the device:** Has this device received 510(k) clearance? What is the "intended use" cleared in the 510(k)?
2. **Population intended for use:** Was the device was intended for a specific population? Expanding its use to other populations may not be assumed; e.g., validation for adults does not confirm that a device is appropriate for newborns.
3. **Conditions or pre-requisites for use:** Are additional tests or procedures to accompany use of the device cited in the FDA clearance?
4. **Ability to use the device exclusively as intended and cleared by FDA:** Are there conditions or restrictions (training requirements, interpretation of results, etc.) which will lead to off-label use and create liability given the circumstances in your institution?

The answers to these questions should be clearly stated in the User Manual for the device and carefully reviewed prior to decision-making. Validated technology that meets regulatory requirements and JCIH recommendations is the critical foundation for successful early hearing detection and intervention. Ineffective or inaccurate technology allows infants to lose the opportunity for identification and optimal outcomes. Failure to detect these infants creates risk and liability for screening programs and the personnel responsible for their administration. By carefully assessing the evidence, however, reliable screening technology can be identified with confidence that it meets the standard of care and that trustworthy documentation supports its selection.

References

1. Durieux-Smith, A, Picton T, Edwards C, Goodman JT, Macmurray B. (1985). The Crib-o-gram in the NICU: an evaluation based on brain stem electrical response audiometry. Ear and Hearing 6:20-24.
2. Fria TJ. (1985). Identification of congenital hearing loss with the auditory brainstem response. In Jacobson JT, ed. The auditory brainstem response. San Diego: College Hill Press, 317-336.
3. Galambos R. Hicks GE, Wilson MJ. (1984). The auditory brainstem response reliably predicts hearing loss in graduates of the tertiary intensive care nursery. Ear and Hearing 5:254-260.
4. Hall JW, Kripal JP, Hepp T. (1988). Newborn hearing screening with auditory brainstem response: measurement problems and solutions. Seminars in Hearing 9:15-32.
5. Herrmann BS, Thornton AR, Joseph JM. (1995). Automated infant hearing screening using the ABR: development and validation. American Journal of Audiology 4 (2):6-14.
6. Hyde ML, Riko K, Corbin H, Moroso M, Alberti PW. (1984). A neonatal hearing screening research program using brainstem electric response audiometry. Journal of Otolaryngology 13(1):49-54.
7. Jacobson JT, Jacobson CA, Spahr RC. (1990). Automated and conventional ABR screening techniques in high-risk infants. Journal of the American Academy of Audiology 1:187-195.
8. Joint Committee on Infant Hearing 2000 Position Statement. http://www.pediatrics.org/cgi/content/full/106/4/798
9. Joint Committee on Infant Hearing 2007 Position Statement Pediatrics 120 (4).
10. Silverman, Franklin H. 1993. Research Design and Evaluation in Speech Pathology and Audiology. Englewood Cliffs NJ: Prentice Hall, 222-233.
11. Schwartz M, Shaw L 1975. Signal Processing: Discrete Spectral Analysis, Detection and Estimation. New York: McGraw-Hill, 216-265.
12. Stein, LK. (1984). Evaluating the efficiency of auditory brainstem response as a neonatal hearing screening test Seminars in Hearing 5:71-76

[1]Clinical and Research Audiologist, original developer of ALGO technology

[2]Executive Director, Audiology & Professional Relations, Natus Medical Incorporated

Healthcare solutions with one thing in mind. You.

natus®

Natus Medical Incorporated

natus.com