Recording Obligatory Cortical Auditory Evoked Potentials in Infants: Quantitative Information on Feasibility and Parent Acceptability

Kevin J. Munro,[1,2] Suzanne C. Purdy,[3] Kai Uus,[1] Anisa Visram,[1,2] Rachel Ward,[2] Iain A. Bruce,[2] Antonia Marsden,[4] Michael A. Stone,[1,2] and Bram Van Dun[5]

Objectives: With the advent of newborn hearing screening and early intervention, there is a growing interest in using supra-threshold obligatory cortical auditory evoked potentials (CAEPs) to complement established pediatric clinical test procedures. The aim of this study was to assess the feasibility, and parent acceptability, of recording infant CAEPs.

Design: Typically developing infants (n = 104) who had passed newborn hearing screening and whose parents expressed no hearing concerns were recruited. Testing was not possible in 6 infants, leaving 98, age range 5 to 39 weeks (mean age = 21.9, SD = 9.4). Three short duration speech-like stimuli (/m/, /g/, /t/) were presented at 65 dB SPL via a loudspeaker at 0° azimuth. Three criteria were used to assess clinical feasibility: (i) median test duration <30 min, (ii) >90% completion rate in a single test session, and (iii) >90% response detection for each stimulus. We also recorded response amplitude, latency, and CAEP signal to noise ratio. Response amplitudes and residual noise levels were compared for Fpz (n = 56) and Cz (n = 42) noninverting electrode locations. Parental acceptability was based on an 8-item questionnaire (7-point scale, 1 being best). In addition, we explored the patient experience in semistructured telephone interviews with seven families.

Results: The median time taken to complete 2 runs for 3 stimuli, including preparation, was 27 min (range 17 to 59 min). Of the 104 infants, 98 (94%) were in an appropriate behavioral state for testing. A further 7 became restless during testing and their results were classified as "inconclusive." In the remaining 91 infants, CAEPs were detected in every case with normal bilateral tympanograms. Detection of CAEPs in response to /m/, /g/, and /t/ in these individuals was 86%, 100%, and 92%, respectively. Residual noise levels and CAEP amplitudes were higher for Cz electrode recordings. Mean scores on the acceptability questionnaire ranged from 1.1 to 2.6. Analysis of interviews indicated that parents found CAEP testing to be a positive experience and recognized the benefit of having an assessment procedure that uses conversational level speech stimuli.

Conclusions: Test duration, completion rates, and response detection rates met (or were close to) our feasibility targets, and parent acceptability was high. CAEPs have the potential to supplement existing practice in 3- to 9-month olds.

Key words: Feasibility, Hearing loss, Infants, Obligatory cortical auditory evoked potentials.

(Ear & Hearing 2020;41;630–639)

INTRODUCTION

In some countries that have implemented a newborn hearing screening program, the age at intervention has dramatically decreased. In England, for example, the median age at which hearing aids are prescribed and fitted has reduced from 300 days of age, when the national program first became fully implemented in 2006 to 2007, to 82 days of age in 2012 to 2013 (Wood et al. 2015). With this reduction in age has come the desire from parents and hearing health professionals to supplement existing clinical practices with new procedures that can be used at 3 to 9 months of age, the period between hearing aid fitting and obtaining reliable behavioral assessment data (primarily via visual reinforcement audiometry [VRA]). In this target age group, there is a need to confirm that speech is being detected or when alternative management strategies should be expedited. This was the motivation for the present study.

Obligatory cortical auditory evoked potentials (CAEPs) are a series of waves, generated in the auditory cortex in response to sound and recorded with surface electrodes on the head. CAEPs can be recorded in awake infants within the first few months of life (Wunderlich et al. 2006). CAEPs show changes in morphology and scalp distribution (Kurtzberg et al. 1984), and amplitude and latency (Ponton et al. 1996; Wunderlich et al. 2006) with maturation. In infants, the CAEP typically consists of a single positive wave around 200 msec after stimulus onset, and a following negativity, rather than the P1–N1–P2 complex recorded in adults. There is some evidence, albeit limited, that the presence of a CAEP correlates with auditory function reported by parents (Golding et al. 2007), and lab-based speech recognition (Rance et al. 2002). Because the CAEP can be evoked by stimuli with a longer duration than the auditory brainstem response (used in most newborn hearing screening programs), it has the potential to verify that conversational-level speech sounds have been physiologically detected at the level of the cortex. The large amplitude and lower temporal precision of the CAEP also make it a useful measure to consider for the assessment and management of infants with auditory neuropathy spectrum disorder (Gardner-Berry et al. 2016).

Rapin and Graziani (1967) were the first to publish aided CAEP data. They demonstrated that it was possible to detect aided CAEPs in children with hearing loss at stimulus levels

[1]Manchester Centre for Audiology and Deafness, School of Health Sciences, University of Manchester, Manchester, United Kingdom; [2]Manchester University Hospitals NHS Foundation Trust, Manchester Academic Health Science Centre, Manchester, United Kingdom; [3]School of Psychology, University of Auckland, Auckland, New Zealand; [4]Centre for Biostatistics, School of Health Sciences, The University of Manchester, Manchester, United Kingdom; and [5]National Acoustic Laboratories, Australian Hearing Hub, Macquarie University, Sydney, Australia.

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and text of this article on the journal's Web site (www.ear-hearing.com).

that could not be detected unaided. Since then, a number of case studies (Gravel et al. 1989; Pearce et al. 2007) have suggested that aided CAEPs may assist in the fitting of hearing aids. This has led to the development of commercially available clinical systems that can be used to record aided CAEPs (Mehta et al. 2017).

Speech-evoked CAEPs have been recorded in children with hearing loss (Purdy et al. 2004). The goal of this approach is to verify if speech, or speech-like sounds, presented at conversational levels, are effectively transduced by the child's hearing aid(s) and detected at the level of the auditory cortex. Note that this technique verifies physiological detection, and presumably audibility, but not that a hearing aid is necessarily meeting prescription targets. The procedure is analogous to the Ling sound test, where a series of isolated phonemes are used to assess low-, middle-, and high-frequency sounds, vital for the development of spoken language skills (Ling 1976; Agung et al. 2005).

Several studies have measured aided CAEPs in infants with hearing loss. Of particular relevance are two relatively small-scale studies that used the same evoked potential system (HEARLab; Frye Electronics, Tigard, OR) and speech stimuli as the present study. Chang et al. (2012) tested 18 infant hearing aid users. However, many of the participants were older (3 to 15 months) than our 3- to 9-month target group. Van Dun et al. (2015) tested 25 infants with hearing loss, some wearing hearing aids. Again, many of the participants were older (8 to 30 months) than our 3- to 9-month target group. It was not possible to detect CAEPs in 30 to 40% and 22 to 28% of participants in Chang et al. and Van Dun et al., respectively, even when the stimuli were expected to be audible. Neither of these studies reported test duration or completion rates. A number of studies have demonstrated that CAEP detection rates increase with sensation level. For example, a retrospective review of clinical data obtained using the HEARLab (Gardner-Berry et al. 2016) demonstrated that CAEP detection increased with sensation level; however, even at sensation levels greater than 20 dB, around 30% of children with sensorineural hearing loss did not have a detectable CAEP.

Despite the availability of clinical evoked potential systems to record infant CAEPs in the sound field, there is a dearth of information on the feasibility and acceptability of recording CAEPs in our target population. Important unanswered questions include, what is the typical test duration, completion rate, and proportion of infants where supra-threshold stimuli can be detected using CAEPs? Based on established pediatric audiology practice test procedures (such as auditory brainstem evoked responses and VRA), we made the assumption that for CAEPs to be adopted in routine clinical practice, the typical test duration would need to be less than 30 min, with successful completion (>90%) and detection (>90%) in most infants. The main aim of the present study was to determine if it was possible to meet these targets in typically developing infants below 9 months of age who passed their newborn hearing screen. We used the opportunity to record response amplitude, latency, and CAEP signal to noise ratio (SNR). The test protocol included the option of placing the noninverting electrode on the forehead instead of the vertex in infants with a hirsute scalp. Although not an original aim, it quickly became apparent that the research audiologists had a preference to use the forehead location instead of the vertex because this avoids the fontanelle and has the perceived benefit that it could reduce preparation time, better electrode retention, and lower contact impedance. Therefore, we incorporated a comparison of the two electrode locations in the data analysis. Because little is known about acceptability of the CAEP test procedure to parents, the second aim of the study was to assess this using a combination of self-report questionnaire and semistructured interviews. The focus here was on how families constructed and framed their experiences of having their infant tested with CAEPs.

MATERIALS AND METHODS

Participants

Ethics approval was granted by the Greater Manchester North National Health Service Research Ethics Committee (Ref 10/H1011/26). A total of 104 infants were recruited, primarily from parent-and-toddler groups. All had passed newborn hearing screening and there was no parental concern about hearing. Families received travel expenses and a £5 voucher for participating in the study. At the start of the test session, parents completed the LittlEARS Auditory Questionnaire that evaluates auditory behavior for children aged up to 24 months (Coninx et al. 2009). All infants who passed tympanometry bilaterally (n = 74) had auditory behavior scores within the LittlEARS 95% reference range for their age. Additional information about the correlation between LittleEARS data and, for example, age, is provided in Appendix 1 in Supplemental Digital Content 1, http://links.lww.com/EANDH/A568.

About half (52%) of the infants were male. Six babies were excluded from the study: 3 drowsy (6, 6, and 12 weeks), 2 unsettled (8 and 15 weeks), and 1 because of a technical error (aged 18 weeks). Removing the one technical error, complete CAEP data were obtained from 98 of 103 (95%) infants. On the day of testing, these 98 infants ranged in age from 5 to 39 weeks (mean = 21.9, SD = 9.4). The majority of infants (93%) were born at 38 to 42 weeks gestational age (mean = 39.9, SD = 1.7); 9 infants were born at 35 to 37 weeks gestational age. Birthweight ranged from 1860 to 4280 g (mean = 3391, SD = 499); only 5 infants weighed less than 2500 g at birth.

Cortical Auditory Evoked Potentials

Testing took place in a single session in the NIHR Manchester Clinical Research Facility at the local National Health Service hospital. The test room (approximately 4 m by 3 m) was not specifically sound treated; rather, it was a quiet room located in a quiet part of the building, away from the main corridor. Before bringing each infant into the test room, the HEARLab control microphone was placed at the approximate location where the infant's head was likely to be located and we measured the ambient noise level (always less than 35 dBA), then completed the HEARLab sound field equalization process and checked the stimulus presentation level. The two testers were research audiologists, with a similar duration of experience in pediatric assessment and electrophysiological testing, and an experienced child playworker. The infant was located 1 m in front of the loudspeaker, usually on the parent's lap, with the child playworker at 45° azimuth. The role of the playworker was to maintain the child in a settled and alert state using a selection of small quiet toys.

CAEPs were recorded to three speech stimuli using the HEARLab (Frye Electronics), a commercially available evoked potential system designed for use with a clinical population. We used this system because of the availability of speech-like stimuli, a calibration and sound field equalization procedure, and the system's automated statistical analysis of the CAEP waveforms. The stimuli, originally extracted from running speech, had spectral power predominantly in the low (/m/), mid (/g/), and high (/t/) frequencies. All stimuli were presented at 65 dB SPL (with an impulse setting on the sound level meter, 35 msec attack-time-constant, slow decay) because these approximately equaled the long-term rms level of the continuous speech from which the stimuli were extracted.

After skin preparation using abrasive tape (3M Red Dot Trace Prep, Maplewood, MN), 3 disposable Ambu Neuroline 720 wet gel self-adhesive electrodes (Ambu, Copenhagen, Denmark) were attached to the scalp. A headband or surgical tape was used to retain the electrodes in position. Electrode impedances were less than 5 kΩ at the start of the recording. Impedances were remeasured at the end of the test session and had increased (to between 6 and 10 kΩ) for 16 (16%) infants. The configuration of the electrodes was: ground electrode on the left mastoid (M1; despite bilateral listening, we selected a single site and for simplicity opted for the same location is all participants), inverting electrode on the right mastoid (M2), and noninverting electrode on the forehead, either just below the hairline (labeled as Fpz) or at the vertex (Cz). The initial protocol specified this vertex location unless the presence of hair made electrode attachment difficult. Completion of the first few infants showed that the pediatric research audiologists preferred the ease of placement at the forehead; therefore, the protocol was modified so that the choice of noninverting electrode location was pseudo-randomized. In total, the Fpz electrode location was used for 56 (57%) infants, including those with hair that was perceived by the tester as potentially impeding good electrode attachment. For the remaining 42 (43%) infants, the noninverting electrode was placed on Cz.

Two runs of a minimum of 75 epochs were recorded for each of the 3 stimuli, presented in blocks of 25 per stimulus, with the stimuli automatically rotating from /m/ to /t/ and then to /g/. The interstimulus interval was 1125 msec. The stop criterion for each stimulus/run was 75 accepted epochs, in complete blocks of 25 presentations. Each run therefore consisted of between 75 and 99 accepted epochs for each stimulus (150 to 198 artifact-free epochs in total for each stimulus). The sampling rate was 1000 samples/sec. Online electroencephalogram (EEG) bandpass filter settings and artifact rejection levels were 0.2 to 30 Hz and ±150 µV, respectively. The tester had the ability to pause acquisition (e.g., if the infant became restless). The first epoch of a test run and the first epoch after a pause were rejected by the HEARLab software. On average, 12.1% (SD = 13.4) of epochs were rejected. During the acquisition of EEG responses, the residual noise was monitored to assess the quality of the averaged CAEP responses. At the completion of CAEP testing, the research audiologists assigned a global rating of overall recording conditions for each infant as good, satisfactory, or poor.

Tympanometry

Tympanometry was performed after CAEP testing was completed using a Titan screening tympanometer (Interacoustics, Middelfart, Denmark). The probe tone was 1000 Hz for babies aged 24 weeks or younger (n = 59) and 226 Hz for older infants (n = 39). Tympanometry was performed on 93 (95%) of the infants with complete CAEP data. There were 75 infants with normal tympanograms (for low-frequency tympanometry: middle ear admittance 0.2 to 1.5 mL, peak middle ear pressure −200 to +50 daPa; for high-frequency tympanometry: positive peak above baseline). Twelve infants were recorded as bilateral fails (11 bilateral flat and 1 flat/negative pressure) and 5 infants were coded as unilateral fail (3 left-ear fails, 2 right-ear fails).

Acceptability of CAEP to Parents

At the end of the test session, parents were informed that they would receive an email in the coming week with a link to an online survey. This was an eight-item questionnaire enquiring about the test procedures, e.g., compared with other tests and procedures that my baby has experienced, tolerating the hearing test was (seven-point scale ranging from “not difficult at all” [1] to “extremely difficult” [7]). A total of 55 parents completed this questionnaire.

Families were invited to be interviewed by telephone a few weeks after testing and we stopped at n = 7 because of data saturation (i.e., no new information after interview five). The interview was conducted by a researcher who was not involved in the CAEP testing and was not aware of the study findings. An exploratory qualitative approach was adopted to explore parents' perceptions of CAEP testing in young babies, and an objective of the research was to use the findings to inform professional guidelines for practice. The semi-structured interview was guided by the following questions: what are your thoughts on having your infant tested with CAEP? How do you think you would feel about the test if your child was hearing/not hearing? What was positive/pleasant about this experience? What was negative/uncomfortable about this experience? How does this experience compare with your other experiences with child health care? and, What advice would you give the professionals who are performing this test? The interviewer clarified points in the narrative as it progressed. The duration of interviews ranged from 7 to 19 min (mean = 11 min, SD = 3). Interviews were conducted in English and were audio-recorded, transcribed, and a thematic content analysis approach was used to interpret the data.

Data Analysis

CAEP Response Detection Using Hotelling's T^2 • CAEP detection was determined using Hotelling's T^2 statistic (Flury & Riedwyl 1988; Golding et al. 2007). Although HEARLab incorporates an automated detection procedure using Hotelling's T^2, for the purposes of this study, we recorded 2 runs of at least 75 epochs and then combined the EEG data off-line into a single file of at least 150 epochs for analysis using MATLAB (MathWorks Inc., version R2015a, MA). The EEG time series of each epoch was averaged across intervals spanning 101 to 500 msec, in increments of 50 msec, such that each 50-msec average formed a 9-point “response.” Hotelling's T^2 calculates the probability that the mean value of any linear combination of the nine variables, in any given time bin, is significantly different from zero. Because of the latency window, Hotelling's captures both P1 and the following negativity in the infant CAEP waveform. Response presence ($p \leq 0.05$) was determined for each stimulus and each child.

CAEP Peak-Picking • Before off-line automatic peak-picking, all data were filtered with a 30 Hz low-pass filter, which matched HEARLab online filtering. P1 amplitudes and latencies were determined using an automatic-peak-picking algorithm implemented in MATLAB followed by a manual check in all cases. P1 was defined as the maximum amplitude in the range 51 to 350 msec. The automatic peak-picking algorithm correctly identified 97% of the peaks. Inaccuracies in the automatic P1 peak-picking algorithm occurred when an identifiable P1 was visible but the automatic peak-picking identified the larger P2 and when P1 had multiple peaks (see example in Appendix 2, Supplemental Digital Content 2, http://links.lww.com/EANDH/A569). After manual checking, 1%, 1%, and 4% of the amplitude or latency values were adjusted for /m/, /g/, and /t/, respectively.

CAEP Waveform SNR • To compare the forehead and vertex locations, SNR analysis of the CAEP waveforms was undertaken off-line in MATLAB for accepted epochs (after artifact rejection at ±150 μV). This off-line analysis used exactly the same code as that utilized by the HEARLab software. SNR was based on CAEP (signal) and residual noise power. The following steps were implemented off-line to calculate SNR:

1. Waveform power was calculated on the averaged waveform in a predetermined range (1 to 600 msec); each sample was squared in the predefined range and the average was obtained.
2. Residual noise power was calculated for each sample (epoch) of the grand average waveform in this predetermined range (1 to 600 msec poststimulus); the variance was calculated around each sample (an indication of noise power per epoch); these were averaged and divided by the number of collected epochs to obtain the residual noise power.
3. SNR based on CAEP (signal) and (residual) noise power: waveform power was corrected by the residual noise power to obtain an (actual) estimate of the CAEP (signal) power. This is based on the assumption that the power in the grand average (the "waveform" power) is the sum of the CAEP (signal) power and the residual noise power. Hence, by subtracting the residual noise power from the waveform power, a more accurate estimate of the power of the CAEP, the signal of interest, is obtained [SNR = 10 × log10 (CAEP power/residual noise power), where CAEP power = waveform power − residual noise power]. Note that negative values occur when the SNR is negative or imaginary, indicating that the CAEP (signal) power is equal to or smaller than the residual noise power, and hence CAEP and residual noise cannot be distinguished. Negative values were therefore converted to zero to avoid skewing the statistics.

This computation of residual noise power assumes that the CAEP is consistent across trials. A limitation inherent to response averaging is this assumption that the morphology of the CAEP is time-invariant. Strait et al. (2014) showed that young children have lower response variability across attentional states compared with older children and adults. The SNR computation is unlikely to be impacted by epoch-to-epoch variations as the SNR in a single trial is very negative and, under these conditions, changing the level of the CAEP response by as much as 100% would have almost no effect on the value of the waveform, relative to the epoch-to-epoch variation due to noise.

TABLE 1. Mean time, in minutes, ± 1 SD for preparation and testing for the 94 infants who completed CAEP data collection and in whom this information was recorded

	Cz (n = 41)	Fpz (n = 53)
Preparation	10.8 ± 5.26	10.6 ± 5.20
Testing	16.1 ± 4.63	15.9 ± 5.30
Total	27.0 ± 7.79	26.5 ± 7.86

CAEP, cortical auditory evoked potential.

Statistical Analyses

Repeated-measures analyses of variance (ANOVA) were performed on P1 latency and amplitude and SNR data using IBM SPSS version 22 software (Chicago, IL). The within-subject factor was stimulus, and between-subject factors were non-inverting electrode location and age group.

RESULTS

Duration of CAEP Test Session

The time, in minutes, for CAEP preparation (instructions, attaching electrodes, and reapplication, where necessary) and CAEP data collection was recorded for 94 of the 98 infants who had complete CAEP data (1 Cz and 3 Fpz not recorded). This is summarized in Table 1. Around 11 min was required for preparation and a further 16 min for data collection. This resulted in overall duration of around 27 min. The differences in time between electrode location were not significant on independent samples *t* tests (preparation $t[92] = 0.26$, $p = 0.80$; data collection $t[92] = 0.18$, $p = 0.90$; total time $t[92] = 0.76$, $p = 0.76$).

CAEP Completion Rate

Within a single test session, CAEP testing was completed in 98 (95%) infants but 7 of these were classified as "inconclusive" because: (i) no or only one CAEP was detected, and (ii) the audiologists described the behavioral state of the infant as "poor."

CAEP Response Detection Using Hotelling's T^2

Table 2 shows the number of infants with responses to the three stimuli. In the 71 infants with normal bilateral tympanograms and acceptable recording conditions, a response was detected to 1, 2, and 3 stimuli in 3%, 20%, and 77% of cases, respectively. The grand average CAEPs, for the infants with bilateral normal tympanograms, are shown in Figure 1. The detection rates decreased in infants with an abnormal tympanogram.

TABLE 2. Number of infants with responses to the three stimuli

	Number of Stimuli Detected			
	0	1	2	3
Bilateral pass (n = 71)	0 (0)	2 (3)	14 (20)	55 (77)
Unilateral fail (n = 9)*	0 (0)	2 (22)	3 (33)	4 (44)
Bilateral fail (n = 11)†	2 (13)	1 (9)	4 (36)	4 (36)

Percentage in parenthesis.
**Includes 4 infants with missing tympanograms in one ear.*
†Includes 3 infants with missing tympanograms bilaterally.

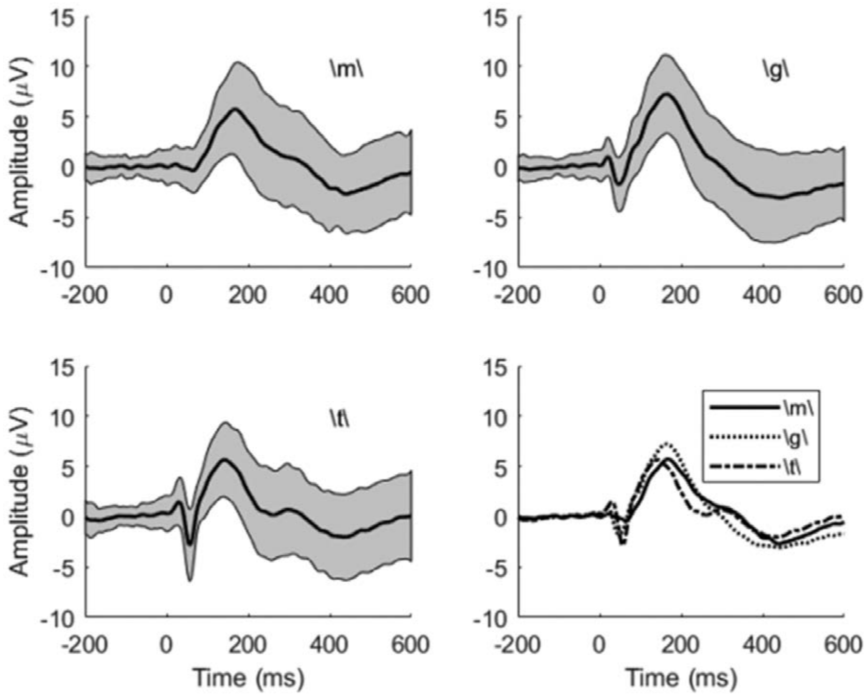

Fig. 1. The grand average CAEP for the 74 participants with normal bilateral tympanograms. Shaded area shows 1 SD around the grand average. CAEP indicates cortical auditory evoked potential.

Table 3 shows response detection rates for each of the three stimuli individually. In infants with normal tympanograms, detection was high for /g/ and /t/. Across all three groups, /m/ had the lowest detection rates.

CAEP SNR at Vertex and Forehead

Figure 2 shows the power of the CAEP (signal) and the residual noise, and the corresponding SNR, at Cz and Fpz. The CAEP power (top panel) was similar for the 2 locations (around 18 μV^2). The residual noise power (middle panel) was around 6 μV^2 at Cz and 4 μV^2 at Fpz. The mean SNR (bottom panel) was similar for the two locations (around 5 dB).

Each of the three variables (CAEP signal, residual noise, and SNR) were analyzed separately using a three-factor ANOVA (between-subject factors: electrode location [2] and age group [2] and within-subject factor is stimulus [3]). The data have been split into two age groups, the cutoff point being set where there were approximately similar numbers in each group.

For CAEP (signal) power, there was a significant difference across stimuli [$F(2, 140) = 3.15$; p=0.046] and no significant interactions with stimuli. The overall effects of age group and electrode location were not significant [$F(1, 70) = 0.289$; $p = 0.593$ and $F(1, 70) = 0.212$; $p = 0.647$, respectively] but there was a significant interaction between age and electrode [$F(1, 70) = 4.84$; p= 0.031]. For residual noise power, electrode location and age group were significant [$F(1,70) = 17.8$; $p < 0.001$ and $F(1,70) = 25.5$; p <0.001, respectively] but there was no difference between the 3 stimuli [$F(2,140) = 0.402$; $p = 0.670$] and no significant interactions. For SNR, the difference at the 2 electrode locations was not significant [$F(1,70) = 0.112$; $p = 0.738$], the difference between age groups was not significant [$F(1,70) = 1.83$; $p = 0.180$], the difference between the stimuli was borderline significant [$F(2,140) = 2.96$; $p = 0.055$] and there were no significant interactions. In summary, CAEP (signal) power and SNR were similar at the two locations, but residual noise power was lower at Fpz. CAEP (signal) power and SNR were similar in the two age groups, but residual noise power was higher in older babies.

TABLE 3. Number of infants with CAEP responses to each of the three stimuli

	/m/	/g/	/t/
Bilateral pass (n = 71)	59 (86)	71 (100)	65 (92)
Unilateral fail (n = 9)*	6 (67)	7 (78)	7 (78)
Bilateral fail (n = 11)†	6 (60)	8 (80)	9 (82)

Percentage in parenthesis.
**Includes 4 infants with missing tympanograms in one ear.*
†Includes 3 infants with missing tympanograms bilaterally.
CAEP, cortical auditory evoked potential.

Latency and Amplitude

Table 4 summarizes the latency and amplitude data. The data have been split into two age groups, the cutoff point being set where there were approximately similar numbers in each group. The most obvious observation is the reduction in latency and amplitude with age (see Fig. 3 for an example). Latency reduces from around 160 to 200 msec in the younger age group to around 135 to 160 msec in the older age group. Amplitude was more variable but reduced from around 8 to 10 μV in the younger age group to nearer 6 to 9 μV in the older age group. The latency and amplitude data were analyzed separately in a three-factor ANOVA (within-subject: stimuli [3]; between-subject age [2]; and electrode location [2]). For latency, there was a significant effect of age and stimuli but not electrode location and no interactions [age: $F(1,67) = 25.6$,

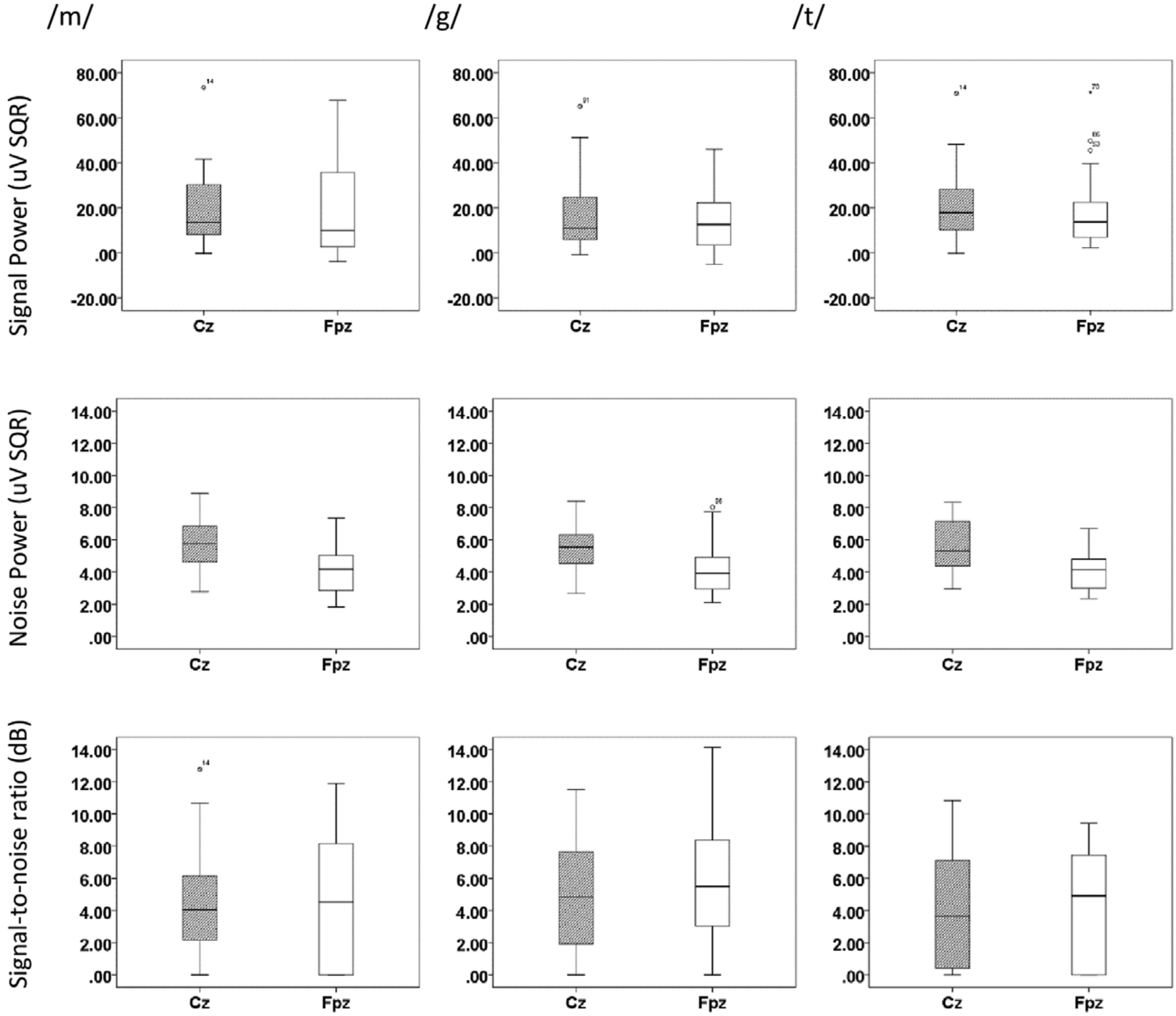

Fig. 2. Comparison of CAEP (signal) and residual noise power (μV^2) and SNR (dB) at Cz (n = 31; filled columns) and Fpz (n = 45; open columns) for the three stimuli for the infants with bilateral normal tympanograms. Top panel: CAEP (signal) power; middle panel: residual noise power; bottom panel: SNR. Note that *y* axis covers a range of 100 and 14 μV^2 for CAEP (signal) and residual noise powers, respectively. CAEP indicates cortical auditory evoked potential; SNR, signal to noise ratio. Small symbols indicate individual outliers.

$p<0.001$; stimuli: $F(2,134)=6.8$, $p=0.02$; electrode location: $F(1,67) = 1.59$, $p = 0.21$]. Posthoc pairwise tests controlling for the average effects of age and electrode location confirmed that the latency of /t/ was shorter than each of the other stimuli ($p < 0.001$ and $p = 0.012$ for the /m/ and /g/ comparisons, respectively). For amplitude, there was a significant effect of age, stimuli, and electrode location but no interactions [age: $F(1,67) = 6.4$, $p = 0.01$; stimuli: $F(2,134) = 5.58$, $p = 0.005$; electrode: $F(1,67) = 7.1$, $p = 0.01$]. Posthoc pairwise tests controlling for the average effects of age and electrode location confirmed that the mean difference between /m/ and /g/ and the mean difference between /g/ and /t/ were significant ($p = 0.002$ for both comparisons).

Parent Acceptability

The responses to the online questionnaire are summarized in Table 5. For almost every question, the typical response was close to the minimum possible response, i.e., very acceptable. The question about keeping their infant "awake but quiet" showed the widest range of responses.

Overall, the interviews revealed that parents were positive about CAEP testing. They discussed the value of the information offered before, during, and after the CAEP recording, the importance of the behavioral state of the infant throughout the CAEP test session, the testers' use of toys to keep the infant happy during testing, and the comfort and accessibility of the test environment. Most parents who discussed receiving information before the test session reported having some awareness of what the appointment would involve. Additional information from the interviews, including quotations from the parents, is provided in Appendix 3 in Supplemental Digital Content 3, http://links.lww.com/EANDH/A570.

TABLE 4. Means and ± 1 SD for P1 latencies and amplitude for the 71 participants with normal tympanograms and interpretable results

	Latency (msec)			Amplitude (μV)		
	/m/	/g/	/t/	/m/	/g/	/t/
Cz						
≤22 mos (n = 14)	186 ± 48	186 ± 23	158 ± 32	10.1 ± 4.2	10.8 ± 4.2	8.1 ± 3.3
>22 mos (n = 15)	144 ± 31	151 ± 23	137 ± 50	6.8 ± 3.7	9.4 ± 2.8	7.9 ± 3.3
Fpz						
≤22 mos (n = 28)	204 ± 39	179 ± 30	172 ± 40	7.6 ± 4.7	8.5 ± 3.9	7.8 ± 3.6
>22 mos (n = 14)	153 ± 20	157 ± 18	146 ± 54	5.6 ± 3.2	6.7 ± 2.8	5.6 ± 2.6

DISCUSSION

The motivation for the study was the desire of hearing professionals and families for clinical procedures that can confirm the detection of speech in infants so that alternative management strategies can be expedited when speech detection is inadequate. We investigated physiological detection of speech sounds using the obligatory CAEPs; specifically, we measured test duration, completion rate, response detection, and parent acceptability.

Duration of CAEP Test Session

The typical test duration for the test paradigm used in the study was 27 min, which was within our 30-min target. There was no difference in mean test time between the two electrode locations. Since the research audiologists decided when to use each location, this might have minimized time differences between electrode locations; for example, preparation may have been longer if using Cz in hirsute infants.

The overall test duration compares favorably with current clinical pediatric test procedures. For example, Janssen et al. (2010) reported that the median time taken to record 6 tone-evoked ABRs thresholds in 72 nonsedated infants (median age, 4 months; 62.5% with normal hearing) was 49 min. None of the parents interviewed in the present study reported that the test session was too long and, in fact, one parent commented that test session was shorter than expected. Therefore, in terms of test duration, these findings suggest that (for the test paradigm used in the present study) CAEPs are feasible for use in the pediatric audiology clinic.

CAEP Completion Rate

Of the 104 infants recruited to the study, 98 (94%) arrived in a settled and appropriate behavioral state for CAEP testing. Of the 98 infants tested, 91 (93%) remained in an appropriate state for testing and we were able to interpret the data. These values exceed our target of 90%. Behavioral state can be an issue for all test procedures in the pediatric audiology clinic. This finding compares favorably with completion rates of existing pediatric clinical procedures. For example, Day et al. (2000) reported that the median number of VRA-obtained minimum response levels at separate frequencies obtained in a single test session was 2.2 for soundfield presentation (n = 22; median age, 32.5 weeks), and 0.9 for insert earphone presentation (n = 19; median age 33 weeks), respectively. This means that multiple test sessions would be required if, for example, the aim was to obtain minimal response levels at four frequencies. Therefore, in terms of completion rate, these findings suggest that (for the test paradigm used in the present study) CAEPs are feasible for use in the pediatric audiology clinic.

CAEP Response Detection

CAEPs were detected in 100% of infants with normal bilateral tympanograms, but not for every stimulus. Because the behavioral thresholds for the stimuli are around 25 dB SPL in adults with normal hearing (Golding et al. 2009), a presentation level of 65 dB SPL corresponds to a sensation level of approximately 40 dB. Detection rates have been shown to increase with sensation level in infants with normal hearing. Cone and Whitaker (2013) reported 100% detection with a stimulus of 60 dB SPL and 85% with a stimulus of 30 dB SPL. Carter et al. (2010) reported a detection rate of 77% with a stimulus sensation level of 30 dB (ca 55 dB SPL) and 54.5% at a sensation level of 20 dB (ca 45 dB SPL). CAEP detection rates for infants with normal hearing (Carter et al. 2010) are not as high as those reported for adults with normal hearing (Golding et al. 2009), presumably because the SNR is degraded in infants due to higher residual background electrophysiological noise.

In the present study, the detection rates varied from 86 to 100%, depending on the stimulus. This was close to our target of >90% for each stimulus. The lowest detection rate was for /m/, which has most energy at the lower frequencies. This stimulus may have been partly masked by vocalizations and other noise from the infant, which will have been predominantly low frequency. Although background noise simultaneous with the stimulus may have led to rejections, noise occurring in the interval before stimulus presentation is likely to have weakened the response without necessarily causing rejection. The rise time of the /m/ stimulus was 10 msec, and that for the /g/ stimulus of 3 msec. The relative broadening of each stimulus due to this gating was therefore similar for the spectral locus of each stimulus. Rise time is therefore not an obvious candidate for the difference observed.

As with all pediatric evoked response testing, when the reliability of the test findings are in question, the results should be treated with caution until reliable data are obtained under good recording conditions.

Overall, CAEP detection rates were lower in the infants with abnormal tympanograms, especially those with bilateral abnormalities. Presumably, this finding is due to reduced audibility resulting in lower amplitude CAEPs. Again, response detection was poorest for the /m/ stimulus. Munro et al. (2011) used the HEARLab to detect CAEPs in young adults using earplugs to simulate a flat conductive hearing loss of around 30 dB HL. In the baseline condition, detection was at, or close to, 100% for each of the three stimuli. In the earplug condition, CAEP

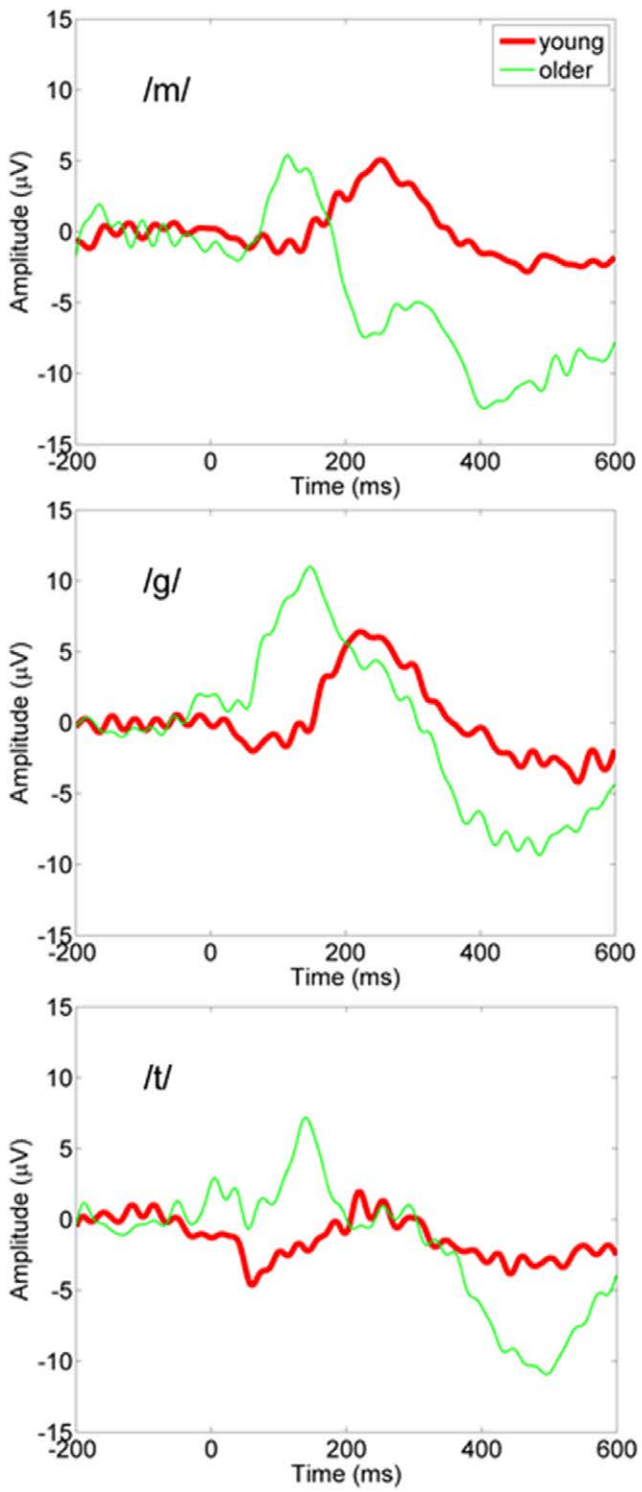

Fig. 3. Comparison of a CAEP response obtained for each of the three stimuli from a younger (bold) and older participant. CAEP indicates cortical auditory evoked potential.

detection reduced, especially at the lowest presentation level (see Table 2 in Munro et al. 2011), consistent with the findings for the infants with abnormal tympanograms in the present study.

The findings from the present study suggest that CAEPs can be reliably detected in infants. However, these data are for infants who passed newborn screening and our primary interest is infants with hearing loss. The clinical utility of CAEPs for determining amplification needs of infants with hearing loss has been reported (Gardner-Berry et al. 2015; Punch et al. 2016; Mehta et al. 2017), but detection rates, and the effects of stimulus type and sensation level, have not yet been examined in detail in a large group of infants in the target age range of 3 to 9 months.

It is important to note, however, that detection of a CAEP does not mean that the sensation level of the amplified stimuli is necessarily optimal for speech and language development (for further information, see Billings et al. 2012). Nor does it mean that audible stimuli can be discriminated. Evoked potential measures that are being used as potential approaches for assessing early infant speech discrimination include the mismatch negativity response and acoustic change complex (see, e.g., Uhler et al. 2018).

CAEP SNR at Vertex and Forehead

Although we originally intended to place the noninverting electrode at Cz for all participants, the research audiologists preferred Fpz for ease of use, especially for infants with a hirsute scalp. There was no significant difference in waveform power between Cz and Fpz (although CAEP peak amplitudes were larger at Cz); however, residual noise power was lower at Fpz and this resulted in a trend for more favorable SNRs at Fpz. The lower residual noise at Fpz is unlikely to be due to systematic differences in physiological noise between the infants in the Cz and Fpz group. However, there may have been difference in nonphysiological noise. For example, it is possible that the electrode at Fpz had a lower impedance (because of less hair and easier-to-abrade skin than Cz) resulting in a more consistent electrode contact over the full test session. Because the bones of the infant skull are thinner and with a space for the anterior fontanelle, this makes it more conductive (Cornelissen et al. 2015). This would suggest, if anything, the power of the CAEP signal would be larger at Cz. For audiologists who are not confident in placing an electrode at Cz, the forehead is a good alternative because skin preparation may be slightly quicker/easier, it avoids the fontanelle, and there is lower residual noise (with no significant difference in SNR between Cz and Fpz).

Latency and Amplitude of the Infant P1

There was a reduction in mean latency and amplitude with age for all stimuli. This finding has been reported many times previously (Wunderlich et al. 2006) and is due to neural maturation. Amplitude is a much more variable measure than latency but is generally larger in infants compared with adults. The peak amplitude of infant P1 can vary from 5 to 15 μV, depending on the stimulus used (Cone & Whitaker 2013). The scalp distribution for N1, the dominant wave in adults, is typically fronto-central, often with a maximum close to Cz (Näätänen & Picton 1987) although Bardy et al. (2015) reported evidence of higher amplitude at Fpz relative to Cz. For this reason, the vertex is usually used as the recording site. It is not clear if the same applies to infants. Sussman et al. (2008) have demonstrated that the scalp distribution for P1, N1, and P2 are all maximal at fronto-central sites in adults, whereas in children, P1 is maximal at frontal sites and P2 is maximal at central sites. In our study, the mean amplitude of P1 was around 2 μV larger

TABLE 5. Mean (SD) and median (minimum to maximum) ratings and scale descriptors for the 8 items in the online questionnaire (n = 55)

	Question	Scale	Mean (SD)	Median (Minimum–Maximum)
1	The information I was given about the hearing test before my baby was tested was	1 = very good 7 = not good at all	1.20 (0.45)	1 (1–3)
2	The test procedure made me feel*	1 = very anxious 7 = not anxious at all	1.07 (0.49)	1 (1–3)
3	During the hearing test, my baby appeared to be	1 = very happy 7 = very unhappy	1.73 (0.99)	1 (1–5)
4	Compared with other tests and procedures that my baby has experienced, tolerating the hearing test seemed*	1 = extremely difficult 7 = not difficult at all	1.58 (0.88)	1 (1–5)
5	Keeping my baby awake and quiet during the hearing test was*	1 = extremely difficult 7 = not difficult at all	2.64 (1.65)	2 (1–7)
6	Seeing the tester attach the recording sensors onto my baby's head made me	1 = not worried at all 7 = extremely worried	1.49 (0.98)	1 (1–5)
7	The information I was provided about the hearing test results after the test session was	1 = very good 7 = not good at all	1.45 (0.88)	1 (1–5)
8	The test environment was	1 = very pleasant 7 = very unpleasant	1.40 (0.69)	1 (1–4)

Scoring reversed before analysis for these items.

at Cz compared with Fpz. Despite no difference in waveform power, a larger amplitude can occur if the P1 is "peaky"; however, this difference in amplitude is small and may not be clinically significant.

CAEP amplitude values at Cz for /m/ and /t/ are similar to those in Purdy et al. (2013) and Golding et al. (2006), but latencies are shorter in the present study. Infants were older in these earlier studies (range 2 to 10, mean 5.3 months in Purdy et al. 2013; range 4 to 7, mean 4.8 months in Golding et al. 2006). Latency differences may reflect peak picking differences due to the automatic algorithm used in the present study versus manual peak detection used by Purdy et al. and Golding et al.

Parent Acceptability

The feedback, both online questionnaire and the subset of interviews from our sample of parents with children who passed newborn hearing screen, was encouraging. There was nothing in the feedback to suggest that infant CAEPs could not be incorporated into the pediatric audiology clinic. However, these were families who had no prior concern about hearing loss and no vested interest in the test procedure or outcome. In qualitative research, the generalizability of results needs to be handled with caution. Our results cannot be translated into acceptability in parents of infants with hearing loss, in routine clinics or in other tests.

CONCLUSIONS

The motivation for the study was the desire of hearing professionals and families for clinical procedures that can confirm detection of speech sounds in infants. We investigated physiological detection of speech sounds using the obligatory CAEPs. It is encouraging that test time, completion rates, and response detection rates met (or were close to) our feasibility targets, and parent acceptability was high. These findings apply to our specific test protocol and for infants who passed the newborn screening. Data are now required from infants who wear hearing aids. CAEPs have the potential to supplement existing practice in 3- to 9-month olds.

ACKNOWLEDGMENTS

Funded by a strategic investment grant from Manchester University Hospitals NHS Foundation Trust. The study was supported by the NIHR Manchester Biomedical Research Centre and the NIHR Manchester Clinical Research Facility. Ruth Nassar and Anne-Marie Dickinson, pediatric research audiologists, recruited the participants and collected the data. Laura McNerlin conducted parent interviews. Michael Maslin contributed to initial data analysis.

K.J.M., S.C.P., K.U., R.B., I.A.B. designed the experiment; A.V. contributed to data collection; K.J.M., S.C.P., A.V., M.A.S., B.V.D., and A.M. analyzed the data; K.J.M., S.C.P., A.V., B.V.D. wrote the article; all authors discussed the results and contributed to revision of the manuscript.

B.V.D. is employed at National Acoustic Laboratories (NAL). The HEARLab is a product developed by NAL, part of a government statutory authority, and the Hearing Cooperative Research Centre (HEARing CRC). At the time of testing, the technology was licensed to Frye Electronics.

Address for correspondence: Kevin J. Munro, Manchester Centre for Audiology and Deafness, University of Manchester, Manchester, M13 9PL, United Kingdom. E-mail: kevin.munro@manchester.ac.uk

Received November 12, 2018; accepted July 15, 2019.

REFERENCES

Agung, K. B., Purdy, S. C., Kitamura, C. (2005). The Ling sound test revisited. *Aust N Z J Audiol*, *27*, 33–41.

Bardy, F., Van Dun, B., Dillon, H. (2015). Bigger is better: Increasing cortical auditory response amplitude via stimulus spectral complexity. *Ear Hear, 36*, 677–687.

Billings, C. J., Papesh, M. A., Penman, T. M., et al. (2012). Clinical use of aided cortical auditory evoked potentials as a measure of physiological detection or physiological discrimination. *Int J Otolaryngol, 2012*, 365752.

Carter, L., Golding, M., Dillon, H., et al. (2010). The detection of infant cortical auditory evoked potentials (CAEPs) using statistical and visual detection techniques. *J Am Acad Audiol, 21*, 347–356.

Chang, H. W., Dillon, H., Carter, L., et al. (2012). The relationship between cortical auditory evoked potential (CAEP) detection and estimated audibility in infants with sensorineural hearing loss. *Int J Audiol, 51*, 663–670.

Cone, B., & Whitaker, R. (2013). Dynamics of infant cortical auditory evoked potentials (CAEPs) for tone and speech tokens. *Int J Pediatr Otorhinolaryngol, 77*, 1162–1173.

Coninx, F., Weichbold, V., Tsiakpini, L., et al. (2009). Validation of the LittlEARS(®) Auditory Questionnaire in children with normal hearing. *Int J Pediatr Otorhinolaryngol, 73*, 1761–1768.

Cornelissen, L., Kim, S. E., Purdon, P. L., et al. (2015). Age-dependent electroencephalogram (EEG) patterns during sevoflurane general anesthesia in infants. *Elife, 4*, e06513.

Day, J., Bamford, J., Parry, G., et al. (2000). Evidence on the efficacy of insert earphone and sound field VRA with young infants. *Br J Audiol, 34*, 329–334.

Flury, B. & Riedwyl, H. (1988). *Multivariate Statistics: A Practical Approach*. London: Chapman and Hall.

Gardner-Berry, K., Purdy, S. C., Ching, T. Y., et al. (2015). The audiological journey and early outcomes of twelve infants with auditory neuropathy spectrum disorder from birth to two years of age. *Int J Audiol, 54*, 524–535.

Gardner-Berry, K., Chang, H., Ching, T. Y., et al. (2016). Detection rates of cortical auditory evoked potentials at different sensation levels in infants with sensory/neural hearing loss and auditory neuropathy spectrum disorder. *Semin Hear, 37*, 53–61.

Golding, M., Purdy, S., Sharma, M., et al. (2006). The effect of stimulus duration and inter-stimulus interval on cortical responses in infants. *Aust N Z J Audiol*, *28*, 122–136.

Golding, M., Pearce, W., Seymour, J., et al. (2007). The relationship between obligatory cortical auditory evoked potentials (CAEPs) and functional measures in young infants. *J Am Acad Audiol, 18*, 117–125.

Golding, M., Dillon, H., Seymour, J., et al. (2009). The detection of adult cortical auditory evoked potentials (CAEPs) using an automated statistic and visual detection. *Int J Audiol, 48*, 833–842.

Gravel, J. S., Kurtzberg, D., Stapells, D. R., *et al*. (1989). Case studies. *Semin Hear*, *10*, 272–287.

Janssen, R. M., Usher, L., Stapells, D. R. (2010). The British Columbia's Children's Hospital tone-evoked auditory brainstem response protocol: How long do infants sleep and how much information can be obtained in one appointment? *Ear Hear, 31*, 722–724.

Kurtzberg, D., Hilpert, P. L., Kreuzer, J. A., et al. (1984). Differential maturation of cortical auditory evoked potentials to speech sounds in normal fullterm and very low-birthweight infants. *Dev Med Child Neurol, 26*, 466–475.

Ling, D. (1976). *Speech and the hearing impaired child: Theory and practice*. *Alexander Graham Bell Association for the Deaf*; Washington, DC.

Litovsky, R. (2015). Development of the auditory system. *Handb Clin Neurol, 129*, 55–72.

Mehta, K., Watkin, P., Baldwin, M., et al. (2017). Role of cortical auditory evoked potentials in reducing the age at hearing aid fitting in children with hearing loss identified by newborn hearing screening. *Trends Hear, 21*, 2331216517744094.

Munro, K. J., Purdy, S. C., Ahmed, S., et al. (2011). Obligatory cortical auditory evoked potential waveform detection and differentiation using a commercially available clinical system: HEARLab™. *Ear Hear, 32*, 782–786.

Näätänen, R., & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure. *Psychophysiology, 24*, 375–425.

Pearce, W., Golding, M., Dillon, H. (2007). Cortical auditory evoked potentials in the assessment of auditory neuropathy: Two case studies. *J Am Acad Audiol, 18*, 380–390.

Ponton, C. W., Don, M., Eggermont, J. J., et al. (1996). Maturation of human cortical auditory function: Differences between normal-hearing children and children with cochlear implants. *Ear Hear, 17*, 430–437.

Punch, S., Van Dun, B., King, A., et al. (2016). Clinical experience of using cortical auditory evoked potentials in the treatment of infant hearing loss in Australia. *Semin Hear, 37*, 36–52.

Purdy, S. C., Katsch, R., Dillon, H.,et al. (2004). Aided cortical auditory evoked potentials for hearing instrument evaluation in infants. In R. C. Sweewald (Ed). A sound foundation through early amplification: Proceedings of 3rd Pediatric Conference, 8, 115–127. Chicago: Phonak AG.

Purdy, S. C., Sharma, M., Munro, K. J., et al. (2013). Stimulus level effects on speech-evoked obligatory cortical auditory evoked potentials in infants with normal hearing. *Clin Neurophysiol, 124*, 474–480.

Rance, G., Cone-Wesson, B., Wunderlich, J., et al. (2002). Speech perception and cortical event related potentials in children with auditory neuropathy. *Ear Hear, 23*, 239–253.

Rapin, I., & Graziani, L. J. (1967). Auditory-evoked responses in normal, brain-damaged, and deaf infants. *Neurology, 17*, 881–894.

Strait, D. L., Slater, J., Abecassis, V., et al. (2014). Cortical response variability as a developmental index of selective auditory attention. *Dev Sci, 17*, 175–186.

Sussman, E., Steinschneider, M., Gumenyuk, V., et al. (2008). The maturation of human evoked brain potentials to sounds presented at different stimulus rates. *Hear Res, 236*, 61–79.

Uhler, K. M., Hunter, S. K., Tierney, E., et al. (2018). The relationship between mismatch response and the acoustic change complex in normal hearing infants. *Clin Neurophysiol, 129*, 1148–1160.

Van Dun, B., Carter, L, Dillon, H. (2015). Sensitivity of cortical auditory evoked potential detection for hearing-impaired infants in response to short speech sounds. *Audiol Res*. *2e13*, 65–76.

Wood, S. A., Sutton, G. J., Davis, A. C. (2015). Performance and characteristics of the Newborn Hearing Screening Programme in England: The first seven years. *Int J Audiol, 54*, 353–358.

Wunderlich, J. L., Cone-Wesson, B. K., Shepherd, R. (2006). Maturation of the cortical auditory evoked potential in infants and young children. *Hear Res, 212*, 185–202.