Published in final edited form as:

Ear Hear. 2013 January ; 34(1): 96–109. doi:10.1097/AUD.0b013e3182650abd.

Determining electrically evoked compound action potential thresholds: A comparison of computer versus human analysis methods

E. Katelyn Glassman, Au.D.^{1,2,3} and Michelle L. Hughes, Ph.D.¹

¹Boys Town National Research Hospital, Omaha, NE

²Division of Speech and Hearing Sciences, University of North Carolina-Chapel Hill, Chapel Hill, NC

Abstract

Objectives—Current cochlear implants (CIs) have telemetry capabilities for measuring the electrically evoked compound action potential (ECAP). Neural Response Telemetry (NRT™; Cochlear) and Neural Response Imaging (NRI; Advanced Bionics [AB]) can measure ECAP responses across a range of stimulus levels to obtain an amplitude growth function. Software-specific algorithms automatically mark the leading negative peak, N1, and the following positive peak/plateau, P2, and apply linear regression to estimate ECAP threshold. Alternatively, clinicians may apply expert judgments to modify the peak markers placed by the software algorithms, and/or use visual detection to identify the lowest level yielding a measurable ECAP response. The goals of this study were to: (1) assess the variability between human and computer decisions for (a) marking N1 and P2, and (b) determination of linear regression threshold (LRT) and visual detection threshold (VDT); and (2) compare LRT and VDT methods within and across human and computer decision methods.

Design—ECAP amplitude growth functions were measured for three electrodes in each of 20 ears (10 Cochlear Nucleus® 24RE/CI512, and 10 AB CII/90K). LRT, defined as the current level yielding an ECAP with zero amplitude, was calculated for both computer- (C-LRT) and human-picked peaks (H-LRT). VDT, defined as the lowest level resulting in a measurable ECAP response, was also calculated for both computer- (C-VDT) and human-picked peaks (H-VDT). Because NRI assigns peak markers to all waveforms but does not include waveforms with amplitudes less than 20 μ V in its regression calculation, C-VDT for AB subjects was defined as the lowest current level yielding an amplitude 20 μ V.

Results—Overall, there were significant correlations between human and computer decisions for peak-marker placement, LRT, and VDT for both manufacturers ($r = 0.78$ to 1.00 , $p < 0.001$). For Cochlear devices, LRT and VDT correlated equally well for both computer- and human-picked peaks ($r = 0.98$ to 0.99 , $p < 0.001$), which likely reflects the well-defined NRT algorithm and the lower noise floor in the 24RE and CI512 devices. For AB devices, correlations between LRT and VDT for both peak-picker methods were weaker than for Cochlear devices ($r = 0.69$ to 0.85 , $p < 0.001$), which likely reflect the higher noise floor of the system. Disagreement between computer and human decisions regarding the presence of an ECAP response occurred for 5.0 % of traces for Cochlear devices and 2.1 % of traces for AB devices.

Corresponding Author: Michelle Hughes; 425 North 30th Street, Omaha, Nebraska 68164; 402-452-5038 (phone); 402-452-5028 (fax); michelle.hughes@boystown.org.

³Current affiliation: New York University Medical Center, New York, NY

Conclusions—Results indicate that human and computer peak-picking methods can be used with similar accuracy for both Cochlear and AB devices. Either C-VDT or C-LRT can be used with equal confidence for Cochlear 24RE and CI512 recipients because both methods are strongly correlated with human decisions. However for AB devices, greater variability exists between different threshold determination methods. This finding should be considered in the context of using ECAP measures to assist with programming CIs.

Keywords

cochlear implant; electrically evoked compound action potential; threshold; Neural Response Telemetry; Neural Response Imaging

INTRODUCTION

Presently, cochlear implant (CI) manufacturers offer devices with telemetry capabilities for measuring the electrically evoked compound action potential (ECAP). Clinically, ECAP measures can provide an objective means to verify neural and implant function by recording the synchronous activity of auditory-nerve fibers in response to electrical stimulation. ECAPs can also be used to monitor physiology over time, guide speech processor fitting when behavioral results are difficult to obtain, and assist in managing complicated cases (e.g., Botros & Psarros, 2010; Franck & Norton, 2001; Han et al., 2005; Holstad et al., 2009; Hughes et al., 2000). The mechanisms for measurement of ECAP responses, known as Neural Response Telemetry for Cochlear devices (NRT™; Cochlear Ltd., Macquarie, New South Wales, Australia) and Neural Response Imaging for Advanced Bionics devices (NRI; Advanced Bionics [AB], Valencia, California, USA), have been incorporated into the respective fitting software to allow for use in device programming. These mechanisms measure ECAP waveforms in response to a range of stimulus levels (i.e., an amplitude growth function) in order to determine the lowest unit of current where an ECAP response is present, known as the visual-detection ECAP threshold.

Several methods exist for determining ECAP thresholds. Clinicians can choose to use thresholds determined by the software or they can use their own judgments in lieu of the software's algorithms. When an ECAP response is measured with the clinical software, the N1 and P2 peaks are automatically picked based on the software's specific algorithms (Fig. 1). In the Cochlear software, two different mechanisms exist for measuring NRT™: an automatic mode, AutoNRT™, and a manual mode, Advanced NRT™. In AutoNRT™, the software starts at a low current level (CL) and increases the level of stimulation until an ECAP measure meeting the algorithm is obtained. Once this ECAP response is measured, the level is decreased to better define the threshold level (Botros et al., 2007). Each waveform is analyzed relative to the waveform collected immediately prior. In Advanced NRT™, the starting level for the stimulus and masker must be obtained prior to obtaining ECAP responses, and each waveform is analyzed as an individual trace (A. Botros personal communication, 2009). For both AutoNRT™ and Advanced NRT™, peaks are picked based on the AutoNRT™ expert system (described in Botros et al., 2007). In this algorithm, N1 is defined as the minimum of the first 8 samples and P2 is the maximum of the samples after N1, up to and including sample 16. When the algorithm is violated, the amplitude of the ECAP response is recorded as 0 µV and peaks are not marked. Possible algorithm violations for Nucleus 24RE or CI512 devices include the following: (1) the negative peak is clipped, (2) the tested electrode is out of voltage compliance, (3) N1 minus P1 is less than 0 µV, (4) the latency between N1 and P2 is less than 2 samples (~98 µsec), and (5) the latency between N1 and P2 is greater than 12 samples (~584 µsec). For AB devices, the NRI peak-picker algorithm uses the first most negative peak value in the range of 270 to 670 µsec and then the following most positive value in the range of 370 µsec or N1 time + 100 µsec

(whichever is greater) to 1000 μ sec to define N1 and P2. The ECAP waveform must follow the end of the stimulus plus 5 samples, and the waveform must be at least 61 samples long ($\sim 18 \mu$ sec sampling interval; P. Trautwein and K. Hood, personal communications, 2009). The software will pick peaks if the tested electrode is out of voltage compliance, but designates those electrodes with a bolded ring around the electrode icon on the software screen. In the NRI algorithm, peaks are picked for all waveforms, but only waveforms with peak-to-peak amplitudes of $\geq 20 \mu$ V are included in the regression analysis that is used to determine ECAP threshold.

Although ECAP N1 and P2 peaks are automatically picked when using the commercial clinical software, a clinician can choose to adjust the peak markers based on his or her own experienced judgments. Using this method, a clinician can choose to include waveforms that violate the software's algorithm or exclude waveforms that meet the parameters of the software's algorithm but appear to be an artifact of the noise floor. Within both computer and human peak-picking methods, ECAP thresholds can be determined through linear-regression analysis or by visual detection.

Software-generated thresholds are presently determined using a linear-regression analysis applied to computer-picked peaks in an amplitude-growth function. These are commonly referred to as T-NRT (Cochlear) and tNRI (AB); in the present study they are referred to as computer linear-regression thresholds (C-LRT). With this method, a regression line is applied to ECAP traces with marked peaks, and threshold is defined as the current level for a zero-amplitude intercept. The linear-regression method requires at least three suprathreshold responses at different current levels to calculate a regression line, which may be difficult to obtain if ECAP thresholds approach the patient's uncomfortable loudness levels (Hughes, 2010). The method assumes that ECAP response amplitudes increase linearly with increases in current level; an assumption that may not always be valid (Botros et al., 2007; Lai & Dillier, 2007). Alternatively, clinicians may use judgments to adjust, add, or delete computer-based peak markers and use the resulting linear-regression threshold based on human-picked peaks. For the present study, this latter threshold-determination method is referred to as human linear-regression threshold (H-LRT).

An alternative method of threshold determination is visual detection. The visual-detection method is more subjective, where threshold is determined as the lowest current level for which an ECAP waveform can be visually identified by a human observer. Unlike the linear regression method which requires multiple responses, the visual detection method utilizes a single response. This method works best in systems with a low noise floor. For systems with a high noise floor, ECAP responses occurring at low current levels may be obscured, resulting in artificially elevated threshold estimates (Hughes, 2010). Although visual detection is a term that traditionally applies to human decisions, visual-detection thresholds (VDTs) can be determined from peak markers placed automatically by computer algorithms, as well as from human-picked peak markers. For this study, these threshold determination methods are referred to as C-VDT and H-VDT, respectively, where C-VDT is defined as the lowest level where the computer algorithm picked peaks that were included in the regression analysis, and H-VDT is defined as the lowest level where a clinician identified an ECAP response based on human-picked peaks.

To date, only a few studies have assessed the variability between different peak-picker and threshold-determination methods. For Cochlear devices, Botros et al. (2007) and van Dijk et al. (2007) compared computer-based VDTs to human-determined VDTs and found less than 9 current-level units (CL^1) difference between the two methods in 90% of cases. In a comparison of VDT and LRT methods, Jeon et al. (2010) found a significant correlation between VDT and LRT ($r = 0.9$) in a group of 12 AB CII/90K recipients; however VDTs

were an average of 23 clinical units (CU²) higher than LRTs. In a group of nine AB CII recipients, Han et al. (2005) found a stronger correlation between most-comfortable (M) levels and ECAP thresholds obtained with visual detection ($r = 0.741$) than with linear regression ($r = 0.675$). Similarly, for adult Cochlear recipients, Potts et al. (2007) found slightly better correlations between comfort (C) levels and ECAP thresholds obtained with visual detection ($r = 0.69$) than with linear regression ($r = 0.66$). In contrast to these findings, somewhat better correlations were found using linear regression than visual detection for pediatric Cochlear recipients (Holstad et al., 2009).

If there are large differences among the methods used to determine ECAP thresholds, clinical interpretations of ECAP responses may be impacted. Automated algorithms are designed to circumvent the variability in the level of expertise across clinicians. However, one issue is that clinicians may not document the specific method used to determine ECAP thresholds in patient reports. The use of different methods among clinicians who follow the same patient can introduce variability into the patient's care. In order for clinicians to use ECAP thresholds to guide programming levels and measure potential changes in physiology over time, it is important to understand how the use of different peak-picker and threshold-determination methods affects ECAP threshold measures. The primary goal of this study was to determine the extent to which ECAP peak markers and thresholds differ for computer algorithms versus human decisions when the same threshold-determination method is used (i.e., linear regression or visual detection). A secondary goal was to examine differences in ECAP thresholds obtained using simple-linear regression analysis versus visual detection.

MATERIALS AND METHODS

Subjects

Three electrodes (one basal, middle, and apical) were tested for each of 20 ears in 20 CI recipients (17 adults, 3 children). Ten subjects were implanted with Cochlear devices (Cochlear Ltd., Macquarie, New South Wales, Australia): nine with the Nucleus[®] 24RE[CA] and one with the Nucleus[®] CI512). Ten subjects were implanted with AB devices (Advanced Bionics, Sylmar, California, USA): five with the CII and five with the HiRes 90K. Subjects were recruited as they were seen for other studies in our laboratory, and had to be at least 11 years old with a minimum of 3 months of device use. Demographic and test-electrode information for participants is provided in Table 1. One Cochlear subject, F10, was originally implanted with an AB HiRes 90K, which was explanted due to device failure after 3 years of use. The subject was reimplanted with a Nucleus 24RE, which was the device tested in this study. For this subject, the age at first implant and total duration of device use is listed in Table 1, with the age at reimplant and duration of 24RE device use listed in parentheses.

Electrode impedance was measured prior to data collection using the clinical-programming software (Custom Sound[™] EP for Cochlear and SoundWave[™] for AB). Electrodes categorized as short or open circuits as determined by the manufacturer's clinical software were excluded from testing. Amplitude growth functions were generally obtained for electrodes 3, 11, and 20 for Cochlear subjects and electrodes 1, 8, and 14 for AB subjects. Deviations from this set were necessary only when electrode impedance was outside the normal range (Subject C19), electrodes were deactivated (Subject C21), or measurable ECAPs could not be recorded (Subject F2). For these subjects, the closest electrode with a

¹CL is a log-based current unit used with Cochlear devices, where current (in microamperes) equals $17.5 \cdot 100^{(CL/255)}$ for Nucleus 24RE or CI512 devices (Botros et al. 2007).

²CU is a linear charge unit used with AB devices, where current (in microamperes) equals $CU/(PW \cdot 0.013)$. The default pulse width (PW) in NRI is 32 microseconds (P. Chellakumar, personal communication, 2010).

measurable ECAP response within the same region (basal, middle, or apical) was tested. For subject C20, ECAPs could not be obtained for any basal electrode (electrodes 12-16; see Table 1). The total number of amplitude growth functions obtained for each device was 30 for Cochlear subjects, and 29 for AB subjects.

ECAP Measures

Cochlear—ECAP thresholds were obtained using the Advanced NRT feature within Custom Sound™ EP (version 2.0.4.7298)³. Stimuli were delivered through a laboratory Freedom speech processor interfaced with a programming pod. Prior to data collection, an ascending procedure was used to estimate behavioral threshold and upper comfort levels for the stimulus used to elicit ECAPs. The masker and probe were set to equal current levels (as described in Abbas et al., 1999) and were increased in steps of 5 CL. The subject was instructed to indicate when the sound was first heard and when the sound was loud but not uncomfortable. These judgments corresponded to ratings of 1 and 8, respectively, on a visual scale of 0–10, where 0 was no sound and 10 was too loud. For ECAP recordings, the masker level was fixed at the current level corresponding to a loudness rating of 8, while the probe level was decreased in 5-CL increments until no response could be visually identified. Additional waveforms were collected in 2-3 CL steps near threshold. The standard forward-masking paradigm (A-B+C-D) was used for artifact reduction (e.g., Abbas et al., 1999, Brown et al., 1998, Dillier et al., 2002).

Recordings were made using the default parameters in the Custom Sound™ EP software (recording electrode located two positions apical to the stimulating electrode). All waveforms were collected at a probe rate of 80 Hz with the masker-probe interval fixed at 400 μ sec. The default gain was 50 dB, with 50 sweeps, and the default sampling delay was 122 μ sec, which was optimized individually as needed. A default pulse width of 25 μ sec/phase was used for both devices except when voltage-compliance limits were reached. This occurred for subject F1, so the pulse width was increased to 50 μ sec to remain within voltage-compliance limits. Recording parameters for all Cochlear subjects were adjusted individually to optimize the ECAP waveform and minimize stimulus artifact (following Abbas et al., 1999). This was typically accomplished by adjusting only the sampling delay. Because the internal chip is the same for 24RE and CI512 devices, both Cochlear device types were analyzed together.

AB—ECAP thresholds were obtained using the NRI feature within SoundWave™ (version 1.4.77). Stimuli were delivered through a laboratory Platinum Series speech processor interfaced with a clinical programming interface (CPI II). ECAPs were obtained using ascending current levels with alternating polarity for artifact reduction (e.g., Miller et al., 1998). The initial stimulation level was set to 50 CU, as this was typically not audible. Current was increased in 15-CU increments until the subject reported that the stimulus reached maximum comfort levels or an ending level of 399 CU was reached, whichever came first. A final level of 399 CU was chosen because it typically remained within voltage-compliance limits and was tolerable by most subjects. ECAP recordings were made using the default parameters of the SoundWave™ software (recording electrode two positions apical to the stimulating electrode). All traces were collected using a cathodic-first sequence with ADC gain of 300 (linear multiplier). ECAPs for both CII and 90K devices were collected using 128 averages with a pulse width of 32 μ sec/phase and a pulse rate of 30 pps. Because the internal chip is the same for CII and HiRes 90K devices, both AB device types were analyzed together. NRI does not allow for the user to adjust many of the measurement

³The Advanced NRT mode was used because it allowed for an analysis of each waveform of the amplitude growth function rather than the threshold alone.

parameters, so if a response could not be obtained by changing the recording electrode, an adjacent stimulating electrode was substituted.

ECAP measures were obtained differently for each manufacturer due to inherent differences between Cochlear and AB software. Cochlear software utilizes forward masking for stimulus artifact reduction. For this method, the masker is fixed at a high level, with amplitude growth functions collected for a descending probe level. Therefore, a loud-but-comfortable starting level had to be obtained using ascending loudness-growth measures for Cochlear subjects prior to ECAP measurements in order to avoid overstimulation. Because AB software only utilizes alternating polarity for artifact reduction, ECAP measurements were obtained in an ascending fashion while simultaneously monitoring loudness growth, similar to the initial loudness measures made for Cochlear subjects. Software usability also affected the method in which ECAPs were obtained for each device. For Cochlear subjects, additional waveforms could easily be added into the existing growth function and incorporated into the LRT function. This capability was not available in the AB SoundWave™ software used for this study. Finally, the difference in current units between manufacturers accounts for the differences in step sizes used for the present study. Cochlear uses log-based CL and AB uses linear units of current (CU). The step size for AB was fixed at 15 CU for the entire growth function, whereas the step size for Cochlear was 5 CL with steps of 2-3 CL near threshold. Because CL is log-based, the equivalent current-unit step size for 2-3 CL varies depending upon the specific value in CL.

Computer Analysis

For the computer-determined measurements (C-LRT and C-VDT), N1 and P2 latencies and peak-to-peak amplitudes were determined for each trace based on the software-specific peak-picker algorithms described in the Introduction. In this phase of the data analysis, there was no human intervention to adjust the peak markers. ECAP thresholds were determined using linear regression and C-VDT (i.e., the lowest level where the computer algorithm identified an ECAP). Simple linear regression was applied to the computer-marked peaks by the automated software algorithms. The linear-regression threshold (C-LRT) was defined by the software algorithms as the current level for a zero-amplitude intercept of the extrapolated regression line. Figure 1 shows screen shots of amplitude growth functions for a Cochlear (top) and an AB (bottom) recipient. For the Cochlear recipient, the Advanced NRT algorithm picked peaks for the waveforms corresponding to 215 CL down to 170 CL (cascade pane on the left), and plotted those peak-to-peak amplitudes in the analysis pane (middle panel on the right). The C-LRT (“intersection”; also called extrapolated T-NRT in the software) was 168.6 CL in this case. For the AB recipient (Fig. 1, bottom), the NRI algorithm picked peaks for all waveforms, and plotted those peak-to-peak amplitudes in the graph on the right. The C-LRT (called tNRI in the software) was 129 CU in this example. It is important to note that in the AB software, peaks are picked for amplitudes $<20 \mu\text{V}$, but those amplitudes are not included in the regression analysis by the algorithm (P. Trautwein personal communication, 2009). Data points that are included in the regression analysis are represented by filled circles, whereas data points not included in the analysis are represented by open circles (the gray circle simply indicates which waveform is highlighted in the left column).

Due to differences in the software peak-picker algorithms, visual-detection thresholds were defined differently between manufacturers. For Cochlear subjects, visual-detection threshold (C-VDT) was defined as the lowest current level where the computer’s algorithm picked peaks. The C-VDT was 170 CL for the Cochlear recipient in Fig. 1 (top). Because the AB NRI algorithm does not include ECAP amplitudes $<20 \mu\text{V}$ in its regression calculation, C-VDT for AB subjects was defined as the lowest current level yielding an amplitude of at least $20 \mu\text{V}$ that was included in the linear-regression analysis. For the AB recipient in Fig.

1 (bottom), the C-VDT was 195 CU (second waveform highlighted in the cascade pane on the left). The ECAP amplitude at 195 CU was 22 μV , whereas the ECAP amplitude for the next lowest stimulus level of 181 CU was 17 μV , which was not included in the regression analysis (top waveform in the cascade pane on the left).

Human Analysis

Two expert clinicians (one was the second author) with 18 years of combined experience working with ECAPs (13 years and 5 years, respectively) examined 1142 ECAP waveforms from 59 electrodes. Once the computer-picked data were stored for separate analysis, peak markers were removed by the first author so that the two clinicians were blinded to threshold, slope, latency, and amplitude data initially determined by the software algorithms. In NRT, peak markers can be manually removed and added. In NRI, peak markers cannot be removed. Therefore, peak markers in the NRI data were manually set to yield peak-to-peak amplitudes of zero, with no regard to latency. Each clinician classified waveforms as “response present” or “response absent” using (1) subjective judgments of the signal-to-noise ratio, (2) overall waveform shape, and (3) peak latencies relative to other traces within the growth function to decide whether identifiable peaks were present in the waveform. Using the method specified by Botros et al. (2007), measurements classified differently by the two clinicians (i.e., present or absent) were disregarded (5.3%). For measurements determined to have an ECAP response, clinicians manually marked N1 and P2 peaks. For measurements determined to have no ECAP response, peak markers were not assigned (Cochlear software) or were adjusted so that peak-to-peak amplitude was zero (AB software).

In order to further examine differences between computer and human peak-picking methods, waveforms judged as no response by either the computer or the clinician were assigned latency values of zero. The reason for this was twofold: (1) the SoundWave™ software does not allow for peak markers to be manually removed, (2) in the Custom Sound™ software, when peaks are left unmarked the amplitude of the ECAP response is recorded as 0 μV with no assigned value for N1 and P2 latency measures. Therefore, in order to determine how often a disagreement existed between the computer and clinicians regarding the presence of an ECAP response, zeros were used as a method for indicating a “response absent” waveform for both manufacturers. Zero data points were used only for latency and amplitude measures as these data were obtained from individual waveforms while threshold data reflected a single data point obtained from each entire amplitude growth function. ECAP thresholds were determined using linear regression and visual detection. Simple linear regression was applied to clinician-picked peaks, and the slope of the line and threshold were calculated as for the computer analysis (H-LRT). Visual-detection threshold (H-VDT) was defined as the lowest current level where each clinician marked N1 and P2 peaks. The inter-judge reliability revealed strong agreement between the two expert clinicians regarding the presence of an ECAP response (94.7%). Therefore for all “response present” measures, the values for N1 and P2 latency, peak-to-peak amplitude, and threshold were averaged across clinicians for comparison with those determined by the software algorithms. Correlations calculated between clinicians were significant across all measures ($r = 0.88$ to 1.00 , $p < 0.001$).

RESULTS

Computer vs. Human: Peak Markers

The first goal of this study was to assess the variability between human and computer decisions for marking ECAP N1 and P2 peaks, and compare linear regression and visual detection methods within and across human and computer threshold determination methods.

In Figure 2, computer- versus human-picked peaks were compared for N1 latency (Figs. 2A-B), P2 latency (Figs. 2C-D), and peak-to-peak amplitude (Figs. 2E-F). Figures in the left column (2A, 2C, and 2E) correspond to Cochlear data and figures in the right column (2B, 2D, and 2F) correspond to AB data. Pearson correlation coefficients were calculated using SigmaStat 3.0 (SPSS Inc., Chicago, IL) for each comparison. Data points included in Figure 2 correspond to waveforms where both computer and human peak-picking methods agreed on the presence of an ECAP response (i.e., without zero data points). The solid black line within each panel of Figure 2 represents the linear regression analysis without the zero data points for Cochlear and AB devices, and the dashed line represents unity. For N1 latency and peak-to-peak amplitude measures, data occurred on both sides of the unity line (i.e., unbiased). For P2 measures, latencies for computer-picked peaks tended to be longer than those for human-picked peaks, particularly for Cochlear devices (Fig. 2C). This may be a result of clinicians picking the P2 peak at the beginning of the plateau, while the software algorithm looks for a latency that falls within a specified time window. A second set of correlations was calculated for *all* measured traces, including those containing zero data points (indicating disagreement between human and computer regarding the presence of a response). Both sets of correlations (with and without zero data points) are reported in Table 2. For these calculations, traces with amplitudes $<20 \mu\text{V}$ were included for AB subjects if the clinicians determined a present ECAP response, but were otherwise excluded because traces with amplitudes $<20 \mu\text{V}$ are not included in NRI linear regression calculations. The mean differences between human- and computer-picked peaks for each device type in Fig. 2 are provided in the first three rows of Table 3. Because differences for individual subjects may be positive (the value for computer-picked peaks is greater than the value for human-picked peaks) while others may be negative (the value for computer-picked peaks is less than the value for human-picked peaks), differences averaged across all subjects were calculated using absolute values. The mean difference between computer- and human-picked peaks was similar between Cochlear and AB devices for N1 latency (2.65 μsec for Cochlear; 5.61 μsec for AB) and amplitude measures (1.04 μV for Cochlear; 1.10 μV for AB), but showed more variability for P2 latencies (34.33 μsec for Cochlear; 17.29 μsec for AB).

Regarding the waveforms for which human and computer decisions agreed on the presence/absence of an ECAP response, Figures 2A and B show significant correlations ($r = 0.98$ for both device groups) between the expert clinicians and software algorithms for N1 peak latency (Table 2). Figures 2C and D show significant correlations ($r = 0.78$ for Cochlear; and $r = 0.95$ for AB) between the expert clinicians and software algorithms for P2 latency. Figures 2E and F show significant correlations ($r = 1.0$ for both device groups) between the expert clinicians and software algorithms for peak-to-peak amplitude.

Regarding the waveforms for which human and computer decisions did not agree on the presence/absence of an ECAP response, correlations were still significant, but not as strong, particularly for the AB device group (Table 2, with zero data points). An example of disagreement between computer and human decisions in AB subjects is shown in Figure 3. Figure 3 shows an example of computer-picked peaks in NRI (top) and human-picked peaks on the same data set (bottom). In the top panel, the software algorithm marked a peak (indicated as the gray bolded circle) that clinicians identified as artifact within the noise floor of the recording system (open bolded circle set to zero in the bottom panel). The gray highlighted peak marked in the top panel corresponds to the trace at 152 CU, which has an amplitude of 21 μV and therefore was included in the computer linear-regression analysis. Traces at 166, 181, 195, and 210 CU, however, were not included in the LRT analysis because the amplitudes are $<20 \mu\text{V}$ and therefore did not meet inclusion criteria. In Fig. 3, C-LRT was 145 CU for the computer-picked peaks (top), and H-LRT was 163 for human-

picked peaks (bottom). Therefore, the inclusion of this one peak in the C-LRT calculation accounts for the large variability between the C-LRT and the H-LRT in this example.

Computer vs. Human: Linear-Regression Thresholds

With the linear-regression method, the regression line can be defined by its slope, and threshold is determined as the x-intercept. For slope, results for both device groups show significant correlations ($r > 0.99$). Correlation coefficients and p values for each manufacturer are reported in Table 4. Figures 4A and B show H-LRTs versus C-LRTs for Cochlear (panel A) and AB subjects (panel B). Results for both device groups show significant correlations ($r > 0.99$ for Cochlear; $r = 0.99$ for AB). Data points were unbiased for both Cochlear and AB devices. Mean differences in both clinical units of measure (CL or CU) and current (μA) are reported in Table 3 for both device groups. For clinical units of measure, mean absolute-value differences were 1.10 CL for Cochlear, and 3.61 CU for AB, which corresponded to mean differences in current of 6.18 μA for Cochlear, and 8.67 μA for AB. On average, the C-LRT was lower than the H-LRT for both devices.

It is important to note that Cochlear and AB devices use different units of measurement for stimulus level (CL and CU, respectively). Because one is logarithmic based (CL) and the other is linear (CU), it is not appropriate to compare correlation coefficients between devices for threshold or slope data. If one wanted to make such a comparison, the data for both devices should be converted to current, in which case the correlation coefficients would change (more so for the log-based CL data). While the primary focus of this study was to compare methods of threshold determination *within* a device type, mean differences and standard deviations in both the stimulus-level units specific to each device and current (μA) are provided in Table 3. This issue is addressed further in the Discussion.

Computer vs. Human: Visual-Detection Thresholds

A common alternative to using C-LRTs (commercially referred to as T-NRT and tNRI) is to visually determine the lowest current level that yields a measurable response (H-VDT). This may not be the same level where the computer picked its lowest current level with a measurable amplitude (C-VDT). Figure 5 shows a comparison of VDTs for human (H-VDT) and computer (C-VDT) peak-picking methods for Cochlear (panel A), and AB (panel B) devices. For Cochlear subjects, C-VDT was defined as the lowest current level where the computer's algorithm picked peaks. Because the AB algorithm does not include ECAP amplitudes $< 20 \mu\text{V}$ (for LRT), visual-detection threshold for AB subjects was defined as the lowest current level yielding an amplitude of $20 \mu\text{V}$. Results for both groups showed significant correlations between human and computer VDTs ($r = 0.98$ for Cochlear; $r = 0.89$ for AB; see Table 4). While data for Cochlear devices were generally unbiased, data from AB devices revealed C-VDTs that were similar to or less than H-VDTs. For clinical units of measure, mean absolute-value differences were 3.25 CL for Cochlear and 10.64 CU for AB, which correspond to mean differences in current of 18.41 μA for Cochlear and 25.58 μA for AB (Table 3). On average, the H-VDT was slightly higher than C-VDT for both devices.

Visual Detection vs. Linear Regression

In order to examine differences that may exist between visual-detection and linear-regression methods, a comparison between VDTs and LRTs was made within each peak-picking method (computer and human). Figure 6 shows a comparison between C-VDTs and C-LRTs for Cochlear (panel A), and AB (panel B) subjects. Results for both device groups, provided in Table 4, show significant correlations ($r = 0.98$ for Cochlear; $r = 0.85$ for AB). H-VDTs and H-LRTs were compared in Figures 6C (Cochlear) and 6D (AB). Results for both device groups show significant correlations ($r = 0.99$ for Cochlear; $r = 0.69$ for AB). For Cochlear devices, data points appeared unbiased for both computer- and human-picked

peaks (Figs. 6A and 6C, respectively). For AB devices, data obtained for computer-picked peaks revealed larger C-VDTs than C-LRTs (Fig. 6B), while data for human-picked peaks (Fig. 6D) showed a greater difference between LRT and VDT than seen for computer-picked peaks. On average there were only slight differences seen between LRT and VDT measures for the Cochlear devices: 4.29 CL (23.36 μ A) for computer-picked peaks and 2.68 CL (14.38 μ A) for human-picked peaks (see Table 3). Mean differences were greater for AB devices, with an average difference of 23.10 CU (55.54 μ A) for computer-picked peaks and 28.54 CU (68.60 μ A) for human-picked peaks. For Cochlear subjects, VDT measures occurred at slightly lower levels than LRT measures, while for AB subjects, VDT measures occurred at higher levels than LRT, as would be expected based on the definition of each measure.

Computer vs. Human: Visual Detection vs. Linear Regression

While visual detection and linear regression can be used with both human and computer peak-picking methods, it may be the case that an inexperienced clinician would use C-LRT, as it requires the least amount of human judgment of the four methods. Likewise, a seasoned clinician who picks peaks based on experienced judgments might elect to define threshold as the lowest level where they can visually detect a response. There would be value in knowing the extent to which the computer-based threshold determinations agreed with those based on visual determination by an experienced human observer. Therefore, visual-detection thresholds using human-picked peaks (H-VDT) were compared with linear-regression analysis of computer-picked peaks (C-LRT). This comparison essentially represents “fully computer” versus “fully human” decisions. Figures 7A-B show a comparison of human VDTs and the computer LRTs obtained with NRT (Cochlear) and NRI (AB), respectively. Results for both device groups, provided in Table 4, showed significant correlations ($r = 0.99$ Cochlear 24RE; $r = 0.71$ for AB). Data obtained for Cochlear devices appeared unbiased. For AB devices, H-VDT data was almost always higher than that obtained for C-LRT. Mean absolute-value differences between C-LRT and H-VDT were 2.70 CL (14.97 μ A) for Cochlear, and 29.57 CU (71.09 μ A) for AB (see Table 3). For both devices, the C-LRT was lower on average than the H-VDT.

DISCUSSION

The primary goal of this study was to determine the extent to which ECAP peak markers and thresholds differ for software algorithms versus human decisions when utilizing linear regression or visual-detection threshold-determination methods. To evaluate these differences, N1 and P2 peaks and ECAP thresholds determined by AB SoundWave™ and Cochlear Ltd. Custom Sound™ EP software algorithms were compared to those determined by two expert clinicians.

Computer vs. Human: Peak Markers

Significant correlations were found between human and computer-picked peaks for N1 and P2 peak latencies and peak-to-peak amplitudes for Cochlear and AB devices (see Fig. 2) both with and without the zero points included in the calculations (Table 2). Results of this analysis revealed low levels of disagreement between human and computer peak-picking methods for both Cochlear and AB devices. Disagreements between human and computer decisions regarding the presence/absence of an ECAP occurred for 5.0% of Cochlear traces, and 2.1% of Advanced Bionics traces. The reason for the large difference between the correlations calculated with the zero data points and without the zero data points varied between the two device groups. For both device types, instances where the *computer* determined an ECAP response to be present but *clinicians* did not only occurred when the software peak-picker algorithm marked peaks for traces that clinicians identified as artifact

within the noise floor of the recording system. Such instances occurred for 32.0% of the disagreements (1.6% of measured waveforms) for Cochlear subjects. For AB subjects, the NRI algorithm in the SoundWave™ software picks peaks for every waveform; therefore instances where the computer picked traces but the clinician did not accounted for 100% of the disagreements. However, when computer-picked peaks with amplitudes $<20 \mu\text{V}$ were excluded from the AB data (because they are not included in AB LRT calculations) the same instance only accounted for 69% of disagreements (1.45% of measured waveforms).

Instances where the *clinicians* determined an ECAP response to be present but the *computer* did not occurred for 68.0% of instances of disagreement (3.4% of measured waveforms) for Cochlear subjects; and 31% (0.65% of measured waveforms) for AB subjects. The reason for this type of disagreement was similar between manufacturers. For both Cochlear and AB subjects, this generally occurred for waveforms in the lower portion of the growth function where the peak-to-peak amplitude of the waveform reached low levels. For Cochlear devices, 88.0% of these instances occurred when the amplitude was $<20 \mu\text{V}$, with the other 12.0% occurring for waveforms with an amplitude between 20 and $30 \mu\text{V}$. For AB subjects, 100.0% of these instances occurred for waveforms where the amplitude was $<20 \mu\text{V}$; therefore, as defined by the NRI algorithm, these waveforms were not included in the regression analysis for C-LRT. For both device types, expert clinicians were able to measure responses at levels lower than $20 \mu\text{V}$, and these waveforms were therefore included in the regression analysis for H-LRT. The overall disagreement for AB subjects (2.1 %) corresponds to instances where the amplitude of the response was at least $20 \mu\text{V}$, or instances where the clinicians measured a response below $20 \mu\text{V}$.

Computer vs. Human: Linear Regression and Visual Detection

For the linear-regression method, comparisons of slope and threshold showed significant correlations between human and computer decisions for Cochlear and AB subjects ($r = 0.99$ to >0.99 ; Table 4, Fig. 4). For slope, correlation coefficients (Table 4) were strong ($r > 0.99$ for both device groups) and mean absolute-value differences (Table 3) were small ($0.34 \mu\text{V}/\text{CL}$ for Cochlear; and $0.02 \mu\text{V}/\text{CU}$ for AB) for all both device groups. In general, the unit of measure used yielded small values for slope for most subjects.

For LRTs, significant correlations were found when comparing human and computer peak-picking methods for both device groups. Likewise, mean absolute-value differences were small for both manufacturers. This suggests that human and computer peak-picking methods can be used with similar reliability for Cochlear and AB devices if LRT is used.

For VDTs, a comparison of human and computer peak-picking methods showed significant correlations for both device groups ($r = 0.89$ to 0.98 ; Table 4, Fig. 5). This means that the lowest current level where the computer's algorithm identified a measurable ECAP was generally consistent with that of the clinician using visual detection with human-picked peaks. Although NRI marks peaks on all waveforms (AB devices), we applied the algorithm's rule for LRT to define C-VDT as the lowest current level that was included in the linear-regression analysis. This generally agreed with H-VDT, and suggests that ECAPs are typically not visually identifiable below approximately $20 \mu\text{V}$ in the AB system. For Cochlear users, Botros et al. (2007) and van Dijk et al. (2007) reported less than 9 CL difference between C-VDT and H-VDT in 90% of cases. A similar comparison is shown in Fig. 5A, where the average (absolute-value) difference between C-VDT and H-VDT for Cochlear subjects was 3.25 CL (range: 0-13 CL; Table 3). There was less than 9 CL difference in 90% of cases for Cochlear devices, which is consistent with that reported by Botros et al. (2007) and van Dijk et al. (2007). For AB data (Fig. 5B), the average difference between C-VDT and H-VDT was 10.64 CU (range: 0-66 CU). Again, note that the correlation coefficients calculated using clinical units of measure cannot be compared across

devices due to differences in the current-level units used (log-based CL versus linear CU). In order to allow for some comparison between devices, clinical units of measure were converted to current in Table 3. In current (μA), mean differences were 18.41 μA (range: 0-82 μA) for Cochlear users, and 25.58 μA (range: 0-158 μA) for AB users. These differences indicate that for some recipients, particularly those with AB devices, greater variability may exist between computer- and human-peak picking methods.

For Cochlear data, there was a significant correlation between LRT and VDT for both computer and human-picked peaks (Fig. 6A and C). On average, there were only slight differences between LRT and VDT measures for computer- and human-picked peaks (4.29 CL and 2.68 CL, respectively; Table 3), and VDT was generally lower than LRT for both computer and human decisions. The correlation between LRT and VDT for human-picked peaks (Fig. 6C) was only slightly stronger than for computer-picked peaks for Cochlear subjects (Fig. 6A).

The consistent results for computer and human decisions for Cochlear devices was likely due to a combination of the well-defined NRT algorithm, which is based on correlations with a large bank of template waveforms (Botros et al., 2007; Van Dijk et al., 2007), as well as the relatively low noise floor for the 24RE and CI512 devices. The noise floor is approximately 2-5 μV for both the Nucleus 24RE and the Nucleus CI512 (Holstad et al., 2009; Patrick et al., 2006; Van Dijk et al., 2007.)

For AB data, the correlation between LRT and VDT was poorer for human-picked peaks (Fig. 6D) than for computer-picked peaks (Fig. 6B). On average, differences between VDT and LRT (absolute value) were 23.1 CU and 28.54 CU for computer and human decisions, respectively (Table 3), and VDT was generally higher than LRT for both human and computer decisions. These results are consistent with those reported by Jeon et al. (2010), who found that VDTs were an average of 23 clinical units (CU) higher than LRTs in a group of 12 AB CII/90K recipients. For computer-picked peaks, the linear-regression function in NRI includes peaks down to 20 μV , and VDT was defined as the lowest level yielding an ECAP amplitude of at least 20 μV . Thus, the decision rules were similar for both C-LRT and C-VDT, leading to good agreement between the two measures. For human-picked peaks, while the correlations were still significant, the decision rules applied to LRT and VDT differed more. In general, VDTs are typically higher than LRTs because of the decision rules associated with each method. Linear regression estimates threshold as the current level for an ECAP amplitude of 0 μV , whereas visual detection requires a measurable ECAP response, which by definition, must be greater than 0 μV . In systems with low noise floors, VDT will more closely approximate LRT, while in systems with high noise floors, there will be more variability. For example, in subjects with noisier responses, the H-VDT may not be visible until the growth function reaches an amplitude much larger than 20 μV , since by definition the H-VDT can only be seen above the noise floor (in NRI, the noise floor is roughly 20-50 μV). In contrast, the C-VDT could theoretically occur within the noise floor, as by definition it only requires the waveform to be included in the linear regression calculation and may or may not correspond to a visible/measurable response. Further, a non-linear growth function can yield a lower LRT than VDT, as it is not restricted by the noise floor. These differences will be accentuated in systems with a relatively high noise floor, where VDT is less accurate.

Perhaps the most clinically applicable comparison is shown in Fig. 7. Those data represent a comparison between strictly automated and strictly human decisions. Inexperienced clinicians may be inclined to rely on computer-picked peaks and LRT, whereas expert clinicians may be inclined to use years of clinical experience to override computer-based peak markers and identify threshold based on visual detection. One disadvantage of the

linear-regression method is that there must be at least three suprathreshold ECAP measures in the amplitude growth function for a regression analysis to be performed. This can be challenging in cases where the ECAP threshold occurs near the recipient's upper loudness tolerance levels (i.e., Hughes, 2010). Further, regression analysis assumes that the function is linear, which in many cases it is not (i.e., Botros et al., 2007). Because of these limitations, VDT is an attractive alternative, particularly in systems with a low noise floor.

If we propose that H-VDT determined by experienced clinicians represents the gold standard, then a comparison of Figs. 5 and 7 within each device can provide useful information regarding the reliability/validity of C-VDT and C-LRT. For Cochlear devices, the correlation coefficients were nearly identical for H-VDT versus C-VDT (Fig. 5A) and H-VDT versus C-LRT (Fig. 7A). This suggests that either C-VDT or C-LRT can be used with equal confidence for Cochlear 24RE and CI512 recipients, and that both methods are significantly correlated with expert human decisions. This finding is in agreement with data from Potts et al. (2007), who reported only slightly better correlations between C-levels and ECAP thresholds obtained with visual detection ($r = 0.69$) than with linear regression ($r = 0.66$) in a group of adult Cochlear recipients.

For AB devices, while the correlations were significant, the correlation between H-VDT and C-VDT was stronger ($r = 0.89$ for AB; Fig. 5B) than for H-VDT versus C-LRT ($r = 0.71$ for AB; Fig. 7B). This suggests that the linear-regression method for AB recipients (i.e., tNRI) is less accurate when compared with expert human decisions.

The comparison between Figs. 5 and 7 shows that the utility of the linear-regression method differs between devices. This is because both manufacturers employ the same analysis (linear regression) on different units of measure for current. For Cochlear data, linear regression is performed on a linear-logarithmic growth function, where ECAP amplitude is a linear unit of measure and current level is logarithmic. In contrast, regression for AB data is performed on a linear-linear function, where both ECAP amplitude and current unit are linear measures. Further, linear-regression analysis assumes that ECAP amplitudes increase linearly with current, which is not always the case (Botros et al., 2007, Lai & Dillier, 2007). On a linear scale, ECAP growth functions seen physiologically tend to be sigmoidal (Miller et al., 2000). This can also happen on a logarithmic scale, which may account for the plateau illustrated in the Cochlear screen capture shown in Fig. 1. In this example, the trace in the middle right-hand panel shows a plateau at 210 and 215 CL. Because these points are included in the linear regression analysis, the slope of the linear regression line is reduced, and the t-NRT (C-LRT) occurs at a lower level than if these points were not included. The plateau in Fig. 1 may have alternatively been caused by saturation of the current source, which would cause a similar effect.

In our clinical observations, it is also not uncommon to see ECAP growth functions with slow growth of amplitude at low levels ("T-tails") in Cochlear recipients, and plateaus at high levels in AB recipients. A log-based measure of current will accentuate the slow-growth portion of the function, making an exponential regression more appropriate than a linear regression, as proposed by Lai and Dillier (2007) for Cochlear users. Although LRT measures correlate well with VDT for most subjects in the present study, in some cases greater variability between these measures could lead to inconsistent applications of ECAP measures for these patients. For example, in Cochlear recipients, a log-based measure of current will de-emphasize plateaus at high levels causing LRT to be overestimated in comparison with VDT. The opposite is true in the case of AB recipients: a linear measure of current will de-emphasize the slow-growth portion at low levels, and emphasize the plateau at high levels causing LRT to occur at a much lower level than VDT. In some instances, the software algorithm may even return negative values for tNRI (C-LRT), and by definition, a

threshold cannot be negative. In any case, it appears that linear regression may not be the most appropriate analysis for automated estimation of ECAP thresholds, regardless of whether linear or logarithmic current measures are used. Thus, it is important to consider (and document) the method used to determine ECAP thresholds when those measures are to be used clinically to assist in creating speech-processor programs or documenting potential changes over time.

Clinically, the relationship between linear regression and visual detection methods of threshold determination will likely affect the accuracy with which ECAP thresholds can be used to assist in the mapping process. For example, Han et al. (2005) found a stronger correlation between comfort (M) levels and ECAP thresholds obtained with visual detection ($r = 0.74$) than with linear regression ($r = 0.68$) in a group of nine AB CII recipients. For Cochlear devices, a study by Potts et al. (2007) found slightly better correlations between comfort (C) levels and ECAP thresholds obtained with visual detection ($r = 0.69$) than with linear regression ($r = 0.66$). While it was not the goal of the present study to compare ECAP threshold determination methods to behavioral map levels (map stimulation rates were not prospectively controlled), it would be a relevant topic for further study.

In conclusion, the results from this study showed significant correlations between computer and expert human decisions within LRT and VDT methods for both Cochlear and AB devices. For Cochlear data, LRT and VDT correlated equally well for both computer-picked and human-picked peaks, which likely reflects the well-defined NRT algorithm and the lower noise floor in the 24RE and CI512 devices. For AB data, the correlation between LRT and VDT was poorer for human-picked peaks than for computer-picked peaks, which is likely due to a combination of differences in the decision rules applied to each method and the higher noise floor of the system.

Finally, for both device groups, the four ECAP threshold determination methods showed significant correlations; therefore clinicians can use any of the four methods to assist in setting map levels. For Cochlear recipients, either C-LRT (ie., T-NRT) or C-VDT can be used with equal confidence because both methods are strongly correlated with expert human decisions. However for AB devices, C-LRT (ie., tNRI) introduces more variability when compared with expert human decisions. Therefore, it is important for clinicians to be aware of the threshold determination method used in order to ensure consistency in programming and document potential changes over time. This finding should be considered in the context of using ECAP measures to program processors for CI recipients.

Acknowledgments

This research was supported by the NIH, NIDCD R01 DC009595, P30 DC04662, and T35 DC008757. The content of this project is solely the responsibility of the authors and does not necessarily represent the official views of the NIH or the NIDCD. The authors thank Lisa Stille (BTNRH) for assistance with data analysis; Jenny Goehring and Jacquelyn Baudhuin (BTNRH) for assistance with data collection; Praveen Chellakumar at Advanced Bionics for technical support with data extraction from the SoundWave™ clinical software; Andrew Botros at Cochlear Ltd. for technical expertise regarding NRT algorithms; and Michael Gorga and three anonymous reviewers for helpful comments on earlier versions of this manuscript.

References

- Abbas PJ, Brown CJ, Shallop JK, et al. Summary of results using the Nucleus CI24M implant to record the electrically evoked compound action potential. *Ear and Hearing*. 1999; 20:45–59. [PubMed: 10037065]
- Botros A, van Dijk B, Killian M. AutoNRT™: An automated system that measures ECAP thresholds with the Nucleus Freedom™ cochlear implant via machine intelligence. *Artificial Intelligence in Medicine*. 2007; 40:15–28. [PubMed: 16920343]

- Botros A, Psarros C. Neural Response Telemetry reconsidered: I. The relevance of ECAP threshold profiles and scaled profiles to cochlear implant fitting. *Ear and Hearing*. 2010; 31:367–379. [PubMed: 20124902]
- Brown CJ, Abbas PJ, Gantz B. Preliminary experience with Neural Response Telemetry in the Nucleus CI24m cochlear implant. *American Journal of Otology*. 1998; 19:320–327. [PubMed: 9596182]
- Dillier N, Lai WK, Almqvist B, et al. Measurement of the electrically evoked compound action potential via a neural response telemetry system. *Annals of Otology, Rhinology & Laryngology*. 2002; 111(5):407–414.
- Franck KH, Norton SJ. Estimation of psychophysical levels using the electrically evoked compound action potential measured with the neural response telemetry capabilities of Cochlear Corporation's CI24M device. *Ear and Hearing*. 2001; 22:289–299. [PubMed: 11527036]
- Han D-M, Chen X-Q, Zhao X-T, et al. Comparisons between Neural Response Imaging thresholds, electrically evoked auditory reflex thresholds and most comfortable loudness levels in CII Bionic Ear users with HiResolution sound processing strategies. *Acta Oto-Laryngologica*. 2005; 125:732–735. [PubMed: 16012035]
- Holstad B, Sonneveldt VG, Fears BT, et al. Relation of electrically evoked compound action potential thresholds to behavioral T- and C-levels in children with cochlear implants. *Ear and Hearing*. 2009; 30:115–127. [PubMed: 19125034]
- Hughes, ML. Fundamentals of clinical ECAP measures in cochlear implants: Part 1: Use of the ECAP in speech processor programming. *Audiology Online*. 2010 Nov 8. Article 2347. Retrieved November 8, 2010 from http://www.audiologyonline.com/articles/article_detail.asp?article_id=2347
- Hughes ML, Brown CJ, Abbas PJ, et al. Comparison of EAP thresholds with MAP levels in the Nucleus 24 cochlear implant: Data from children. *Ear and Hearing*. 2000; 21:164–174. [PubMed: 10777023]
- Jeon EK, Brown CJ, Etler CP, et al. Comparison of electrically evoked compound action potential thresholds and loudness estimates for the stimuli used to program the Advanced Bionics cochlear implant. *Journal of the American Academy of Audiology*. 2010; 21:16–27. [PubMed: 20085196]
- Kwon BJ, van den Honert C. Effect of electrode configuration on psychophysical forward masking in cochlear implant listeners. *Journal of the Acoustical Society of America*. 2006; 119(5):2994–3002. [PubMed: 16708955]
- Lai WK, Dillier N. Comparing neural response telemetry amplitude growth functions with loudness growth functions: Preliminary results. *Ear and Hearing*. 2007; 28:42S–45S. [PubMed: 17496644]
- McKay C, Fewster L, Dawson P. A different approach to using neural response telemetry for automated cochlear implant processor programming. *Ear and Hearing*. 2005; 26:38S–44S. [PubMed: 16082266]
- Miller CA, Abbas PJ, Brown CJ. An improved method of reducing stimulus artifact in the electrically evoked whole-nerve potential. *Ear & Hearing*. 2000; 21:280–290. [PubMed: 10981604]
- Miller CA, Abbas PJ, Rubinstein JT, Robinson BK, Matsuoka AJ, Woodworth G. Electrically evoked compound action potentials of guinea pig and cat: responses to monopolar, monophasic stimulation. *Hearing Research*. 1998; 119:142–154. [PubMed: 9641327]
- Patrick JF, Busby PA, Gibson PJ. The development of the Nucleus® Freedom™ cochlear implant system. *Trends in Amplification*. 2006; 10(4):175–200. [PubMed: 17172547]
- Potts LG, Skinner MW, Gotter BD, et al. Relation between Neural Response Telemetry thresholds, T- and C-levels, and loudness judgments in 12 adult Nucleus 24 cochlear implant recipients. *Ear and Hearing*. 2007; 28:495–511. [PubMed: 17609612]
- van Dijk B, Botros A, Battmer R-D, et al. Clinical results of AutoNRT™, a completely automatic ECAP recording system for cochlear implants. *Ear and Hearing*. 2007; 28:558–570. [PubMed: 17609616]

Figure 1.

Screen captures of amplitude growth functions obtained with the clinical software for a Cochlear (top panel: Custom Sound™ EP) and an Advanced Bionics (AB; bottom panel: SoundWave™) recipient. Waveforms are shown at left for each device, with N1 and P2 peaks marked by vertical lines for Cochlear and circles for AB. Clinical units of measurement (CL for Cochlear, CU for AB) are listed to the right of each waveform. For AB, the amplitude of the response is listed as “EP” in this same location. The right panels for each manufacturer show the linear regression calculations. For AB, the shaded circles indicate waveforms that have been included in the regression analysis. The zero intercept of the calculated line (tNRI = 129) is provided at the top right of this panel. The gray shaded circle corresponds to the currently selected waveform, (195 CU), which in this example represents the C-VDT. For Cochlear, the regression analysis is shown with the slope and the intersection (T-NRT, 168.5975).

Figure 2.

Comparison of computer vs. human picked-peaks for N1 latency (panels A-B), P2 latency (panels C-D), and peak-to-peak amplitude (panels E-F). Cochlear data are represented by open circles in the left column and Advanced Bionics data are represented by closed triangles in the right column. Data points contained in this figure represent only instances where the computer and expert clinicians were in agreement regarding the presence of an ECAP response. Pearson correlation coefficients and statistical significance for the figure are displayed in Table 2. In each panel the solid line represents linear regression analysis and the dashed line represents unity. Data points falling above unity indicate that the computer picked longer latencies than humans, while data points falling below the line indicate the computer picked shorter latencies.

Figure 3. Individual example of a single amplitude growth function obtained with Neural Response Imaging (NRI) for an Advanced Bionics subject. Top: computer-picked peaks; bottom: human-picked peaks for the same data set. In the top panel, the software algorithm marked peaks (open circles in the graph on the right represent ECAP amplitudes less than $20 \mu\text{V}$) that clinicians identified as artifact within the noise floor of the recording system (amplitudes set to zero in the bottom right panel). The gray bolded circle in the top panel corresponds to the selected trace in the left panel. This trace shows a waveform at 152 CU which has an amplitude of $21 \mu\text{V}$ (listed as EP in the software). Because the amplitude of this trace is greater than $20 \mu\text{V}$, it has been included in the linear regression analysis. In the bottom panel, this same trace (bolded circle) has an amplitude of $0 \mu\text{V}$, indicating that the clinicians identified the waveform as artifact, and therefore it was not included in the linear regression analysis.

Figure 4.

Comparison of linear regression thresholds (LRT) for human (H-LRT) and computer (C-LRT) peak-picking methods. Cochlear data are represented by open circles in the left panel and Advanced Bionics data are represented by closed triangles in the right panel. In each panel the solid line represents linear regression analysis and the dashed line represents unity. Pearson correlation coefficients and statistical significance for the figure are displayed in Table 4.

Figure 5.

Comparison of visual detection thresholds (VDT) for human (H-VDT) and computer (C-VDT) peak-picking methods. Left panel: Cochlear, right panel: Advanced Bionics. In each panel the solid line represents linear regression analysis for Cochlear and AB devices, and the dashed line represents unity. Pearson correlation coefficients and statistical significance for the figure are displayed in Table 4.

Figure 6.

Comparison between linear regression thresholds (LRT) and visual detection thresholds (VDT) within each peak-picking method (top: computer, bottom: human) for Cochlear (left column) and Advanced Bionics (right column). Solid lines represent linear regression analyses, and dashed lines represent unity. Pearson correlation coefficients and statistical significance for the figure are displayed in Table 4.

Figure 7.

Comparison of human expert decisions (H-VDT) and fully automated decisions (C-LRT) for Cochlear (panel A) and Advanced Bionics (panel B) subjects. Solid lines represent linear regression analyses, and dashed lines represent unity. Pearson correlation coefficients and statistical significance for the figure are displayed in Table 4.

Table 1

Subject demographics and electrodes tested.

Subject	Device Type	Ear	Age At CI	Duration of CI Use	Apical Elec.	Middle Elec.	Basal Elec.
F1*	Nuc 24RE(CA)	L	60y 7m	3y 2m	20	11	3
F2	Nuc 24RE(CA)	R	60y 2m	2y 1m	20	11	5
F4	Nuc 24RE(CA)	L	17y 6m	1y 11m	20	11	3
F5*	Nuc 24RE(CA)	R	48y 3m	0y 11m	20	11	3
F7	Nuc 24RE(CA)	R	39y 1m	2y 3m	20	11	3
F8*	Nuc 24RE(CA)	L	58y 3m	0y 4m	20	11	3
F9	Nuc 24RE(CA)	R	61y 1m	2y 10m	20	11	3
F10*	Nuc 24RE(CA)	R	8y 3m (11y 1m) [^]	6y 1m (3y 3m) [^]	19	11	3
F15*	Nuc 24RE(CA)	L	22y 9m	2y 7m	20	8	3
N5	Nuc CI512(CA)	R	50y 8m	1y 3m	20	11	3
C1	AB CII	R	18y 4m	7y 0m	1	8	14
C6	AB CII	L	64y 4m	6y 0m	1	8	14
C8	AB CII	R	55y 7m	6y 0m	1	8	14
C13	AB CII	R	77y 1m	6y 6m	1	8	14
C14	AB CII	L	6y 2m	7y 10m	1	8	14
C15	AB HiRes 90K	L	39y 3m	4y 9m	1	8	14
C17	AB HiRes 90K	R	7y 11m	3y 10m	1	8	14
C19	AB HiRes 90K	R	15y 5m	5y 3m	1	8	15
C20	AB HiRes 90K	R	56y 4m	4y 0m	1	8	14**
C21	AB HiRes 90K	R	47y 7m	0y 9m	1	8	11
Mean:							

* Subject is bilaterally implanted. Only the ear listed was tested.

[^] Subject initially implanted with a HiRes 90K, which failed and was replaced with a 24RE. Age at implant and duration of CI use for the device tested is listed in parentheses.

** No measurable ECAP response for electrodes 12-16.

Table 2

Pearson correlation coefficients (r-values) and significance (p-values) for data shown in Figure 2.

Computer vs. Human Peak-picking methods	Without zero data points		With zero data points	
	r	p	r	p
A: N1 latency (Cochlear)	0.98	<0.001	0.86	<0.001
B: N1 latency (AB)	0.98	<0.001	0.49	<0.001
C: P2 latency (Cochlear)	0.78	<0.001	0.87	<0.001
D: P2 latency (AB)	0.95	<0.001	0.46	<0.001
E: Amplitude (Cochlear)	1.00	<0.001	1.00	<0.001
F: Amplitude (AB)	1.00	<0.001	1.00	<0.001

Table 3

Summary of mean differences (absolute value) in latency, peak-to-peak amplitude, slope of the regression line, and ECAP threshold per device type.

Computer-Human Differences	Cochlear	AB
N1 latency (Fig. 2)	2.65 μ sec (9.55 SD)	5.61 μ sec (10.22 SD)
P2 latency (Fig. 2)	34.33 μ sec (64.37 SD)	17.29 μ sec (30.58 SD)
P-P amplitude (Fig. 2)	1.04 μ V (2.91 SD)	1.10 μ V (1.39 SD)
LR slope (not shown)	0.34 μ V/CL (0.42 SD)	0.02 μ V/CU (0.04 SD)

Threshold Differences	CL	μ A	CU	μ A
CLRT-HLRT (Fig. 4C-D)	1.10 CL (0.99 SD)	6.18 μ A (6.16 SD)	3.61 CU (5.31 SD)	8.67 μ A (12.76 SD)
CVDT-HVDT (Fig. 5)	3.25 CL (3.32 SD)	18.41 μ A (20.07 SD)	10.64 CU (17.76 SD)	25.58 μ A (42.70 SD)
CLRT-CVDT (Fig. 6A-B)	4.29 CL (3.35 SD)	23.36 μ A (21.11 SD)	23.10 CU (24.94 SD)	55.54 μ A (59.96 SD)
HLRT-HVDT (Fig. 6C-D)	2.68 CL (2.06 SD)	14.38 μ A (12.11 SD)	28.54 CU (30.13 SD)	68.60 μ A (72.43 SD)
CLRT-HVDT (Fig. 7)	2.70 CL (2.57 SD)	14.97 μ A (15.00 SD)	29.57 CU (28.52 SD)	71.09 μ A (68.55 SD)

SD = 1 standard deviation; P-P = peak-to-peak; LR = linear regression; CLRT = computer linear regression threshold; HLRT = human linear regression threshold; CVDT = computer visual detection threshold; HVDT = human visual detection threshold.

Table 4

Pearson correlation coefficients (r-values) and significance (p-values) for slope and threshold data shown in Figures 4 to 7.

	Cochlear		AB	
	r	p	r	p
Computer-Human LR slope	>0.99	<0.001	>0.99	<0.001
CLRT-HLRT (Fig. 4)	>0.99	<0.001	0.99	<0.001
CVDT-HVDT (Fig. 5)	0.98	<0.001	0.89	<0.001
CLRT-CVDT (Fig. 6A-B)	0.98	<0.001	0.85	<0.001
HLRT-HVDT (Fig. 6C-D)	0.99	<0.001	0.69	<0.001
CLRT-HVDT (Fig. 7)	0.99	<0.001	0.71	<0.001

LR = linear regression; CLRT = computer linear regression threshold; HLRT = human linear regression threshold; CVDT = computer visual detection threshold; HVDT = human visual detection threshold. R-values indicated as >0.99 were between 1.0 and 0.99.