



Published in final edited form as:

Ann Neurol. 2021 April ; 89(4): 790–802. doi:10.1002/ana.26029.

Multisite Study of Evoked Potentials in Rett Syndrome

Joni N. Saby, PhD¹, Timothy A. Benke, MD, PhD², Sarika U. Peters, PhD³, Shannon M. Standridge, MD⁴, Junko Matsuzaki, PhD¹, Clare Cutri-French, BA⁵, Lindsay C. Swanson, MS CGC⁶, David N. Lieberman, MD PhD⁶, Alexandra P. Key, PhD⁷, Alan K. Percy, MD⁸, Jeffrey L. Neul, MD PhD³, Charles A. Nelson, PhD⁹, Timothy P.L. Roberts, PhD¹, Eric D. Marsh, MD, PhD⁵

¹Division of Radiology Research, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania

²Department of Pediatrics, Neurology, Pharmacology and Otolaryngology, University of Colorado School of Medicine and Children's Hospital Colorado, Aurora, Colorado

³Department of Pediatrics, Vanderbilt University Medical Center, Vanderbilt Kennedy Center, Nashville, Tennessee

⁴Cincinnati Children's Hospital Medical Center, Division of Neurology and University of Cincinnati College of Medicine, Cincinnati, Ohio

⁵Division of Child Neurology, Children's Hospital of Philadelphia, Departments of Neurology and Pediatrics, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, Pennsylvania

⁶Department of Neurology, Boston Children's Hospital, Boston, Massachusetts

⁷Department of Hearing and Speech Sciences, Vanderbilt University Medical Center, Vanderbilt Kennedy Center, Nashville, Tennessee

⁸Department of Pediatrics, University of Alabama at Birmingham, Birmingham, Alabama

⁹Laboratories of Cognitive Neuroscience, Boston Children's Hospital, Boston, Massachusetts; Department of Pediatrics, Harvard Medical School

Abstract

Objective: The aim of the current study was to evaluate the utility of evoked potentials as a biomarker of cortical function in Rett syndrome (RTT). As a number of disease-modifying therapeutics are currently under development, there is a pressing need for biomarkers to objectively and precisely assess the effectiveness of these treatments.

Method: Yearly visual (VEP) and auditory (AEP) evoked potentials were acquired from individuals with RTT, aged 2 to 37 years, and control participants across five sites as part of

Corresponding Author: Eric D. Marsh, MD, Division of Child Neurology, Abramson Research Building- Room 502E, 3615 Civic Center Boulevard, Philadelphia, PA, 19104, Phone- 215-590-5654, Fax- 215-590-3779, marshe@email.chop.edu.

Author Contributions:

TAB, AKP, JLN, CAN, TPLR, and EDM contributed to the conception and design of the study. JNS, TAB, SUP, SMS, JM, CC-F, LCS, DNL, AK, JLN, CAN, TPLR, and EDM contributed to the acquisition and analysis of data. JNS, TAB, SUP, AKP, JLN, CAN, TPLR, and EDM contributed to drafting the text and preparing the figures.

Potential Conflicts of Interest: All authors have nothing to report

the Rett Syndrome and Related Disorders Natural History Study. Baseline and Year 1 data, when available, were analyzed and the repeatability of the results was tested. Two syndrome-specific measures from the Natural History Study were used for evaluating the clinical relevance of the VEP and AEP parameters.

Results: At the baseline study, group level comparisons revealed reduced VEP and AEP amplitude in RTT compared to control participants. Further analyses within the RTT group indicated that this reduction was associated with RTT-related symptoms, with greater severity associated with lower VEP and AEP amplitude. In participants with RTT, VEP and AEP amplitude was also negatively associated with age. Year 1 follow-up data analyses yielded similar findings and evidence of repeatability of EPs at the individual level.

Interpretation: The present findings indicate the promise of EPs as an objective measure of disease severity in individuals with RTT. Our multisite approach demonstrates potential research and clinical applications to provide unbiased assessment of disease staging, prognosis, and response to therapy.

Rett syndrome (RTT) is an X-linked neurodevelopmental disorder associated with pathogenic variants in the methyl-CpG-binding protein 2 (*MECP2*) gene [1, 2]. RTT affects predominantly females and is characterized by developmental regression, loss of speech and purposeful hand use, onset of stereotypic hand movements and gait abnormalities [3].

Currently, no disease modifying treatments exist for RTT. Fortunately, many new therapeutics are being developed and tested at the preclinical level [4–7] and a few have made it to clinical trials (e.g. [NCT04181723](#), [NCT03758924](#)). With novel treatments on the horizon, the identification and validation of central nervous system biomarkers in RTT will be useful to assess objectively both early treatment responsiveness, and ultimately, therapeutic efficacy.

The present study seeks to determine the potential utility of evoked potentials (EPs) as a clinically relevant biomarker of cortical function in RTT. Previous studies of EPs in RTT have reported altered latencies and amplitudes in this participant group, albeit with a high degree of interindividual variability [8–15]. Most of these studies had small sample sizes and therefore, lacked sufficient power to examine the potential clinical relevance of the observed variability in EP responses. Recent work within the visual domain provided initial evidence for a link between EP abnormalities and disease severity in RTT [12], but this analysis was performed at the group rather than individual level.

For the current study, visual (VEP) and auditory (AEP) evoked potentials were acquired from a representative cohort of individuals with RTT as part of the multi-site Natural History Study (NHS) of Rett syndrome and Related Disorders. All participants contributing EPs were simultaneously enrolled in the clinical portion of the NHS allowing for comparison of VEP and AEP parameters with clinically relevant phenotypic data, including symptom severity. For a subset of participants, EP and clinical data was available at two time points, which allowed us to examine the repeatability of the results and intersession reliability of the EPs at the individual level.

Methods

Participants

Participants were enrolled in the NHS of Rett and Related Disorders ([NCT00299312/NCT02738281](#)) protocols 5211 and 5212. Evoked potentials were acquired across five study locations: Boston Children's Hospital (BCH), University of Colorado/Children's Hospital Colorado (UC-CHCO), Children's Hospital of Philadelphia (CHOP), Cincinnati Children's Hospital (CCH), and Vanderbilt University Medical Center (VUMC). The experimental protocol was approved by the Institutional Review Boards at each site. For the natural history protocol (5211) the appropriate Institutional Review Boards of CHOP and VUMC approved the protocol, whereas UC-CHCO, BCH, and CCH relied on the single-IRB agreement provided by the University of Alabama at Birmingham. Written informed consent was obtained for all participants/legal guardian.

Baseline evoked potentials were acquired from 70 participants with RTT. Final analyses of the Baseline EPs are based on 41 participants for the VEP and 47 participants for the AEP. The other participants were excluded for excessive artifact or other factors (see Figure 1 for exclusion reasons). Of the 41 participants included in the Baseline analysis of the VEP, Year 1 VEPs (acquired 10 – 14 months later) were available in 24 participants and evaluable in 17. Of the 47 participants in the Baseline analysis of the AEP, Year 1 AEPs were available in 24 and evaluable in 21 (see Figure 1). All participants with RTT had a pathogenic variant in the MECP2 gene identified and confirmed diagnosis of classic, atypical preserved speech, or atypical unspecified RTT (see Supplementary Table 1 for detailed demographics). Diagnosis was based on the published clinical criteria [3] via direct clinical evaluation by an experienced child neurologist or pediatric geneticist.

Twenty-four typically developing (TD) participants (16 females, 8 males; median = 11.0 years of age; range = 2.1 – 25.4 years) with no history of developmental delay or known neurologic, neuropsychiatric, or genetic condition served as the control group. Four additional TD participants were excluded for excessive artifact ($n = 1$) or absence of a clear N1 in the AEP ($n = 3$). Data acquisition for control participants was distributed across study sites. Potential TD participants were screened using the Child/Adult Behavior Checklist [CBCL; 16] and the Ages and Stages Questionnaire [ASQ-3; 17] or Wide Range Achievement Test-4 [WRAT-4; 18] for participants less than or greater than 5 years respectively. TD subjects were included only if they scored within the normal range on the CBCL and within the average or above score for the ASQ-3/WRAT-4. Year 1 EPs were only available in three TD participants and therefore, group comparisons were limited to analyses of the Baseline data. There were no differences in the age of the RTT vs. TD subjects that were included in the VEP ($U = 431.5$, $p = .411$) or AEP ($U = 542.0$, $p = .789$) analyses.

Measures of Clinical Severity

Two clinician-completed measures were used to estimate RTT severity: Clinical Severity Score (CSS; [19, 20]) and Motor-Behavioral Assessment (MBA; [20]). The CSS has 13 items with a maximum score of 58. The MBA has 34 items with a maximum score of 136. CSS and MBA scores for the study participants are provided in Figure 1 (and Supplementary

Table 1). Clinical scores were typically obtained on the same day as the EPs. In the case the visits did not occur on the same day, the scores from the closest clinical visit were used for correlating with the EP parameters.

EP Recording

All sites followed standardized procedures for the VEP and AEP recording. Participants were seated 60 cm from a monitor in a dimly lit room. The visual stimuli consisted of 400 trials of a reversing black and white checkerboard presented continuously (0.5 cpd, 100% contrast, 2 Hz refresh rate). At one site (BCH), pattern reversal was contingent on the participant attending to the stimulus. Participant's gaze was registered using a Tobii T120 Eye Tracker (Tobii Technology, Danderyd, Sweden) and the experiment halted until gaze was redirected at the screen for a minimum of 100 ms. For all sites, a research assistant and/or parent was with the child to redirect their attention when necessary. The auditory stimuli consisted of 520 trials of 500 Hz sinusoidal tones (300 ms duration) presented using a free-field speaker at 60dB SPL. The interstimulus interval was jittered between 0.6 and 2 seconds. A silent movie was presented on the monitor during the AEP task to keep participants calm and still. Both stimuli were presented using ePrime software (Psychology Software Tools, Pittsburgh, PA). The order of tasks was randomized between participants.

EEG Methods

EEG acquisition equipment varied by site.—At CCH, EEG was recorded from 21 individual Ag/AgCl electrodes referenced to FPz and amplified using a Natus EEG32U Amplifier (Natus Neuro, Middleton, WI) digitized at 512 Hz. At CHOP, EEG was acquired from a 60-channel Ag/AgCl electrode cap (EASYCAP GmbH, Herrsching-Breitbrunn, Germany) with a FCz reference. Signals were amplified using the EEG amplifier of an Elekta VectorView System (Elekta-Neuromag, Helsinki, Finland) digitized at 1000 Hz. For the other sites, a 128-channel Geodesic (VMC and UC-CHCO) or HydroCel (BCH) Sensor Net (Electrical Geodesics, Inc, Eugene, OR, USA) was used, referenced to Cz, and digitized at 1000 Hz. At all sites, EEG was acquired continuously with a 0.1 Hz high-pass and 100 Hz low-pass filter. Electrode impedances were checked before all recordings and kept below the individual systems' recommendations.

Prior to initiating data acquisition, a traveling adult human phantom completed the VEP/AEP at all locations. The phantom data revealed similar EP waveform morphology at all sites but with variable timing and amplitude of the components (see Figure 2). To account for these inter-site differences, the human phantom data was used to generate latency and amplitude adjustment factors that were applied to all RTT and TD participants data prior to final analysis. The adjustment factors were site specific, determined based on the difference between each site and CCH. CCH was chosen as the reference site for having the largest amplitudes of the five study locations. VEP data were adjusted based on the difference between P1 latency and amplitude at each site and P1 latency and amplitude at CCH. AEP data were adjusted based on the difference between N1 latency and amplitude at each site and N1 latency and amplitude at CCH. To ensure the adjustment did not alter the results, the analyses (described below) were also conducted on the non-adjusted data. For group-level comparisons, the non-adjusted analysis was conducted using an ANOVA

(equal variances not assumed) with Site as an analysis factor. For associations with clinical severity, the non-adjusted analysis was conducted using multiple regression with Site coded as a dummy variable. The results of these additional analyses were consistent with the main results reported below (see Supplementary Data).

Data analysis was performed at one central location (CHOP) using BESA software (BESA 6.0 GmbH, Grafelfing, Germany). Files collected at 1,000 Hz were first downsampled to match the lowest sampling rate at acquisition (512 Hz). The files were visually inspected for bad channels and periods of excessive artifact, which were marked and excluded from further analysis. Ocular artifacts were removed using automatic artifact correction methods in BESA. The artifact-corrected data were transformed to a reference-free, 81-channel array using spherical spline interpolation [21] to allow for cross-site comparisons. Additional analyses conducted using the original EEG arrays for approximately 1/3 subjects ensured that re-montaging to 81-channels did not distort the latency and amplitude estimates at the channels of interest (Oz and FCz). After re-montaging, data files were digitally filtered using a 3 Hz high-pass filter to avoid the potential confounding influence of low-frequency activity (which was prevalent in the EEG of many of the RTT participants) on EP amplitude estimates. Filtered data were segmented into 500 ms epochs (for the VEP; -100 to 400ms relative to stimulus onset) or 600 ms (for the AEP; -150 to 450 ms relative to stimulus onset) and baseline corrected based on the mean of the pre-stimulus period. Prior to averaging, trials were low-pass filtered at 40 Hz and excluded if the amplitude exceeded $\pm 250 \mu\text{V}$ at one or more channels at any point during the 500 ms (VEP) or 600 ms (AEP) epoch. Due to more pervasive artifact in the EEG files for participants with RTT, the number of accepted trials was 9.7% lower for the RTT than the TD group for the VEP (RTT: median = 342 (range = 162 – 399); TD: median = 377, range = 288 – 400; $U = 267.5$, $p = 0.002$) and 9.2% lower for the RTT than TD group for the AEP (RTT: median = 466, range = 294 – 519); TD: median = 511, range = 440 – 519; $U = 253.0$, $p < 0.001$).

Analysis

Analysis of the VEP focused on the N1, P1, and N2 components of the response at the midline occipital electrode (Oz). In one RTT participant, the right occipital electrode (O2) was used due to the absence of a response at O1 and Oz. The P1 was defined as the first positive component closest to 100 ms. The N1 was defined as the negative peak immediately preceding the P1 and the N2 as the negative peak immediately following the P1. Analysis of the AEP focused on the P1, N1, and P2 components of the response at the frontal-central midline electrode (FCz). The N1 was defined as the first negative peak closest to 100 ms and the P1 and P2 were defined as the positive peaks immediately preceding and following the N1. Peak latencies and amplitudes were identified and measured using the peak finder in BESA. Interpeak amplitudes were measured peak to trough.

Statistics.

Statistical analyses were conducted in IBM SPSS Statistics 26 (IBM Corp, Armonk, NY). Latency and amplitude of the VEP and AEP components were compared between the RTT and TD groups using non-parametric Mann-Whitney U tests due to non-normality and unequal variances between groups. Linear regression analysis examined potential

associations between age and VEP/AEP parameters for the RTT and TD groups and within the RTT group, associations between VEP/AEP parameters and clinical severity (CSS and MBA). When a particular VEP/AEP parameter was associated with both age and clinical severity, a hierarchical regression model was used to determine the individual contribution of clinical severity on the VEP/AEP parameter after accounting for the effect of age. Age was entered at the first step (Step 1) of the model. Clinical severity (CSS or MBA) was entered at the second and final step (Step 2). The log transformation (\log_{10}) of age was entered in all analyses using age to account for a positive skew in the age of the participants. Results were considered statistically significant at $p < 0.05$.

The agreement between Baseline and Year 1 EPs were characterized using intraclass correlation coefficients (ICC). ICC estimates were computed using a two-way mixed effects model with absolute agreement, single measure. Bland-Altman plots (Bland & Altman, 1986) were created to visualize differences in the Baseline vs. Year 1 EPs and to identify subjects with more pronounced changes in EP latency and/or amplitude. ICCs and Bland-Altman plots were computed for absolute, rather than interpeak, amplitude in order to better delineate stability and/or change in each component.

Results

Visual Evoked Potentials: Baseline Data

Comparison of RTT and TD groups.—The grand average VEP waveforms for the RTT and TD groups are displayed in Figure 3A. Participants with RTT exhibited significantly reduced N1 ($U = 229.0, p < 0.001$), N1–P1 ($U = 170.0, p < 0.001$), and P1–N2 ($U = 199.0, p < 0.001$) amplitudes compared to the TD group (Figure 3B). No differences in the timing of the VEP components between the RTT and TD groups existed (N1 latency: $U = 380.5, p = .129$; P1 latency: $U = 354.5, p = .061$; N2 latency: $U = 395.5, p = .189$; P1–N2 time: $U = 441.5, p = .492$).

Associations between VEP and Age.—There was a significant association between \log_{10} age and N1–P1 ($p = 0.015$) and P1–N2 ($p = 0.020$) amplitude for the RTT group, with a decline in amplitude with increasing age (Figure 3C). Age was not significantly associated with the latency or amplitude of any of the VEP parameters for TD participants (Table 1), as expected based on prior research with TD children [22, 23].

Associations between VEP and Clinical Severity.—Within the RTT group, clinical severity (measured by either the CSS or MBA) was associated with N1 (CSS: $p = 0.022$; MBA: $p = 0.035$) and N1–P1 (CSS: $p = 0.010$; MBA: $p = 0.008$) amplitude, with greater severity predicting lower amplitude (Figure 5A). No significant associations were found between disease severity and the latency of the VEP components (Table 1). Given N1–P1 amplitude was also associated with age, a hierarchical regression analysis was conducted to examine the independent contribution of clinical severity on N1–P1 amplitude accounting for the effect of age. The results of this analysis indicated that clinical severity measured by the CSS and MBA accounted for a significant proportion of the variance in N1–P1 amplitude over and above the variance accounted for by age (Figure 5B). Specifically, the addition of CSS to the model with age accounted for an additional 9.6% of the variance in

N1–P1 amplitude and this change in R^2 was significant ($R^2 = .096$, $F(1, 38) = 4.785$, $p = 0.035$). For each unitary change in CSS, N1–P1 amplitude decreased by $.069 \mu\text{V}$ [95% CI = $-.133, -.005$]. The addition of MBA to the model with age accounted for an additional 8.7% of the variance in N1–P1 amplitude. This change in R^2 was also significant ($R^2 = .087$, $F(1, 38) = 4.294$, $p = 0.045$). For each unitary change in MBA, N1–P1 amplitude decreased by $.044 \mu\text{V}$ [95% CI = $-.087, -.001$].

Auditory Evoked Potentials: Baseline Data

Comparison of RTT and TD groups.—The grand average AEP waveforms for the RTT and TD groups are displayed in Figure 4A. Participants with RTT exhibited significantly lower P1–N1 ($U = 386.0$, $p = 0.030$) and N1–P2 ($U = 351.0$, $p = 0.010$) amplitudes compared to TD subjects. There were no group differences in the amplitude of the initial P1 component ($U = 491.0$, $p = 0.375$) or in the latency of any of the AEP components (P1 latency: $U = 501.0$, $p = 0.444$; N1 latency: $U = 527.0$, $p = 0.653$; P2 latency: $U = 492.5$, $p = 0.385$; Figure 4B).

Associations between AEP and Age.— \log_{10} age was significantly associated with the latency of the P1, N1, and P2 components for both RTT ($ps < 0.05$) and TD participants ($ps < 0.001$) with a decline in latency with increasing age. Within the RTT group, \log_{10} age was also negatively associated with P1 ($p = 0.003$) and P1–N1 ($p = 0.001$) amplitude (Figure 4C). No association was observed between \log_{10} age and AEP amplitude for the TD group (Table 2).

Associations between AEP and Clinical Severity.—Clinical severity was significantly associated with AEP amplitude, with greater severity associated with lower amplitude. CSS was negatively associated with P1 ($p = 0.004$) and P1 – N1 ($p = 0.007$; Figure 5C) amplitude and MBA was negatively associated with P1 amplitude ($p = 0.033$). No associations were found between clinical severity and the latency of any of the AEP components (Table 2). Hierarchical regression analyses indicated that CSS accounted for a significant proportion of the variance in P1 ($R^2 = .116$, $F(1, 44) = 7.264$, $p = 0.010$) and P1–N1 amplitude ($R^2 = .095$, $F(1, 44) = 5.996$, $p = 0.018$) after controlling for the effect of age (Figure 5D). For each change in CSS, P1 amplitude decreased by $.027 \mu\text{V}$ [95% CI = $-.047, -.007$] and P1–N1 amplitude decreased by $.044 \mu\text{V}$ [95% CI = $-.080, -.008$]. MBA, in contrast, was no longer a significant predictor of P1 amplitude after accounting for age ($R^2 = .047$, $F(1, 44) = 2.682$, $p = 0.109$).

Visual Evoked Potentials: Year 1 Data

Associations between VEP and Clinical Severity.—The above analyses were repeated using the clinical and EP data from the Year 1 visit. Initial analyses with all participants with evaluable Year 1 VEPs ($n = 17$) did not yield any significant associations between the VEPs and clinical severity. After removing two cases that were identified as influential (defined as Cook's $D > 1$), N1–P1 amplitude was associated with CSS ($p = 0.046$; Figure 5E) as well as participant age ($p = 0.034$). The association between CSS and N1–P1 amplitude was no longer significant when also accounting for the effect of age ($R^2 = .124$, $F(1, 12) = 2.584$, $p = 0.134$; Figure 5F).

Baseline/Year 1 Comparison.—Intraclass correlation analysis indicated strong agreement in severity scores between Baseline and Year 1 (ICCs>.8; Figure 6A). The mean change from Baseline to Year 1 was -0.41 points for the CSS and -0.47 points for the MBA. The latency of the VEP components also demonstrated strong consistency between the two visits (ICCs = 0.75; Figure 6C). The amplitudes of the VEPs were more variable, with poor intersession agreement for N1 amplitude (ICC=0.030) and moderate agreement for P1 amplitude (ICC=0.419) and N2 amplitude (ICC=0.500).

For select subjects, there was strong intersession agreement in latency and amplitude of the components (see Figure 6B). In order to identify whether the variability observed for other subjects was related to changes in clinical severity, we first identified subjects whose score on the CSS and/or MBA changed more than 1 SD of the mean from Baseline to Year 1. We then used Bland-Altman plots to visualize whether participants with greater change in clinical score had a greater change in VEP latency and/or amplitude. A consistent pattern of change in any of the VEP parameters was not tied to change in clinical score (Figure 6). Specifically, a number of subjects who demonstrated more notable change in the VEP, particularly in VEP amplitude, had stable clinical scores. Furthermore, when subjects with clinical change did demonstrate change in the VEP, the direction of the change was not systematically tied to the direction of change in clinical score (improvement vs. worsening of symptoms).

Auditory Evoked Potentials: Year 1 Data

Associations between AEP and Clinical Severity.—Analysis of the Year 1 data confirmed the significant association between AEP amplitude and clinical severity found at baseline, with greater severity predicting lower amplitude. P1–N1 amplitude was associated with both clinical measures (CSS: $p=0.047$; MBA: $p=0.047$; see Figure 5G). The association between MBA and P1–N1 amplitude remained significant when entered into a hierarchical regression model accounting for the effect of age ($R^2=0.197$, $F(1, 18) = 4.488$, $p=0.048$). The association between CSS and P1–N1 amplitude was no longer significant after accounting for the effect of age, but just above the 0.05 significance threshold ($R^2=0.184$, $F(1, 18) = 4.114$, $p=0.058$). P1 amplitude was also associated with clinical severity as measured by the MBA ($p=0.038$) and this association remained significant after accounting for age ($R^2 = .234$, $F(1, 18) = 5.557$, $p=0.030$; Figure 5H).

Baseline/Year 1 Comparison.—Intraclass correlation analysis indicated that CSS and MBA scores remained consistent between Baseline and Year 1 for participants included in the analysis of Year 1 AEPs (ICCs>.9). The ICCs for the AEPs revealed poor inter-session agreement in P1 latency and P1 amplitude (ICCs<.500), but moderate agreement for the latency and amplitude of the N1 and P2 components (ICCs = .500; see Figure 6C).

There was strong consistency in the latency and amplitude of the AEP components for a number of subjects (see Figure 6E). Bland-Altman plots were created to determine if the variability observed for other subjects was related to change in clinical severity. Similar to the results of the VEP analysis, no consistent pattern of change in the AEP related to change in clinical score.

Discussion

This is the first multisite study of EPs in individuals with RTT. Group-level analyses demonstrated attenuation of visual and auditory EPs in RTT compared to TD participants. The finding of attenuated responses in participants with RTT is consistent with prior single-site studies of EPs in RTT [8, 10, 12]. Additionally, VEP and AEP responses demonstrated significant associations to clinical RTT severity scores within individuals. Greater severity on the CSS and the MBA was associated with lower amplitude of the EP components. For both the VEP and AEP, the association between amplitude with clinical severity was specific to the initial peak and the interpeak amplitude between the initial peak and the subsequent peak (N1–P1 for VEP; P1–N1 for AEP). A similar association was not found between clinical severity and amplitude of the latter peak of the waveform (P1–N2 for VEP; N1–P2 for AEP), possibly because these peaks are more variable between RTT participants and often harder to define.

Within individuals with RTT, VEP and AEP amplitude was also negatively associated with participant age. This reduction in response amplitude may reflect a generalized decline in neurologic function with age. Although RTT is not a degenerative condition, the severity of RTT-related symptoms tends to increase with age. The finding that VEP and AEP amplitude decline with age in RTT suggests that these measures may reflect a progressive aspect of the disease process in addition to correlating with inter-individual differences in disease severity.

For a subset of participants, EP and clinical data was available from a follow-up visit occurring one year after the initial recording. Analyses of the Year 1 EPs revealed an association between EP amplitude and clinical severity, validating the results of the Baseline analysis. In the Year 1 data, the association between clinical severity and response amplitude was stronger for the AEP, perhaps due to the relatively larger sample available for the AEP analysis.

Additional analyses of the Year 1 EPs indicated a good degree of intersession repeatability of the EPs for many participants. The finding that the waveforms remained consistent across time in individuals with RTT is promising for the future use of EPs as biomarkers in this population. There were also a number of participants, however, who failed to demonstrate repeatability in aspects of the VEP and/or AEP. The intersession variability for these subjects could not be accounted for by changes in clinical score. The absence of an association between change in clinical score and change in the EPs is somewhat unexpected given the present finding of an association between clinical severity and EP amplitude. The absence of an association between these factors may be attributed to two sources of variability: 1) Variability in clinical scores arising from clinician variability and differences in behavior/state of the subject and 2) Variability in the EPs arising from differences in behavior (e.g., those that result in EEG artifact) and participant arousal/attention to the stimulus. While there was stability in clinical scores from Baseline to Year 1, the repeatability of the CSS and MBA has not been demonstrated, and some variability in scoring from visit to visit in an otherwise stable patient occurs. This point recapitulates the need for a biomarker of severity in RTT that may be more stable than behavior-based questionnaires. Although there was variability in the EPs as well, the finding of stable EPs

in some patients is promising that this variability can be reduced, perhaps by ensuring all behavioral and technical factors remain constant between visits.

This study is the first to report a correlation between EPs and clinical severity in participants with RTT. With continued progress in developing targeted treatments for RTT [4, 7, 24], identifying and validating sensitive biomarkers to objectively test the effectiveness of treatments is paramount. The finding that visual and auditory EPs are attenuated in RTT and that the extent of this change is related to clinical severity, points to the potential utility of EP amplitude as a biomarker of brain function in this group. Research at the preclinical level has reported similar abnormalities in the VEP and AEP in animal models of the condition [12, 25–28] and has identified potential mechanisms underlying the attenuation of the responses [for reviews, see 29, 30]. The translatable value of these measures further points to their potential utility in clinical trials for RTT.

A number of limitations of the current study warrant follow up in future work. Eye tracking was only available at one of the study locations resulting in a difference in when the stimuli were presented; continuously vs gaze fixation required (at the one site). Since attention affects VEP amplitude [31, 32], the potential influence of attention on the current results cannot be ruled out. However, our findings are closely aligned with those of Leblanc and colleagues [12], who utilized eye tracking with all participants. This similarity in findings suggest that direct attention (as assessed by eye tracking) may not be as essential as presumed, at least in neurologically impaired populations such as individuals with RTT. Nonetheless, future studies should employ eye tracking to rule out potential effects of attention. A measure of visual acuity in each participant would also be valuable for accounting for influences of visual acuity on the VEP [11]. Furthermore, data acquired from the traveling human phantom indicated a degree of cross-site variability in the latency and amplitude of the EPs, which was reproduced in the data files for the RTT and control participants. Given RTT is rare, large studies to identify and validate biological markers will undoubtedly require multisite research. As this area of research moves towards within-subject analyses (e.g., pre- vs. post-treatment), differences across sites will not be as consequential. However, for studies that will require averaging across sites, the stimulus presentation and acquisition methods (including electrode type and impedances) should idealistically be matched at all locations to minimize variability.

In order for EPs to be useful in the context of clinical trials, future work needs to 1) determine if EPs in RTT are responsive to treatment, 2) establish the criteria for which EPs in RTT are most reproducible, 3) further characterize how EPs in RTT change overtime in a given participant and 4) determine which component/combination of components is most suitable as a biomarker. Given EEG data from participants with RTT is often contaminated by artifact from movement, bruxism, and other behaviors, this area of research would additionally benefit from the optimization of acquisition and analysis methods for reducing data loss (as was an issue in this study). For the present analyses, a number of RTT participants had to be excluded because their EP waveform did not conform to the expected pattern based on the waveform observed in neurotypical individuals. Although the use of conventional methods allows comparisons to the larger EP literature, the application of non-traditional analysis approaches such as template matching or frequency analysis may

be more inclusive and perhaps more precise for characterizing EPs in RTT. Relatedly, more precise interpretation of EPs in RTT would also require a better understanding of how unique features of the EEG in this population, such as predominant low frequency activity and epileptiform abnormalities, interact with the EP measures.

In summary, despite the limitations, the current finding of an association between VEP and AEP amplitude and RTT severity in a multisite study indicates the potential utility of EPs as a biological marker of neural function in RTT and underscores the need for additional research in this area.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements:

The authors would like to thank all of the subjects and their parents for contributing to this study. We would also like to thank each site's technicians and research assistants who helped with arranging studies and running the protocols. We would also like to thank NIH/NICHHD for funding (U54 HD061222 Project 8880) and RettSyndrome.org for coordination of funding and publicity of our study.

References

1. Amir RE, Van den Veyver IB, Wan M et al. Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat Genet* 1999; 23: 185–188. [PubMed: 10508514]
2. Leonard H, Bower C, English D. The prevalence and incidence of Rett syndrome in Australia. *Eur Child Adolesc Psychiatry* 1997; 6 Suppl 1: 8–10. [PubMed: 9452912]
3. Neul JL, Kaufmann WE, Glaze DG et al. Rett syndrome: revised diagnostic criteria and nomenclature. *Ann Neurol* 2010; 68: 944–950. [PubMed: 21154482]
4. Katz DM, Bird A, Coenraads M et al. Rett Syndrome: Crossing the Threshold to Clinical Translation. *Trends Neurosci* 2016; 39: 100–113. [PubMed: 26830113]
5. Leonard H, Cobb S, Downs J. Clinical and biological progress over 50 years in Rett syndrome. *Nat Rev Neurol* 2017; 13: 37–51. [PubMed: 27934853]
6. Percy AK. Progress in Rett Syndrome: from discovery to clinical trials. *Wien Med Wochenschr* 2016; 166: 325–332. [PubMed: 27491553]
7. Pozzo-Miller L, Pati S, Percy AK. Rett Syndrome: Reaching for Clinical Trials. *Neurotherapeutics* 2015; 12: 631–640. [PubMed: 25861995]
8. Foxe JJ, Burke KM, Andrade GN et al. Automatic cortical representation of auditory pitch changes in Rett syndrome. *J Neurodev Disord* 2016; 8: 34. [PubMed: 27594924]
9. Bader GG, Witt-Engerstrom I, Hagberg B. Neurophysiological findings in the Rett syndrome, II: Visual and auditory brainstem, middle and late evoked responses. *Brain Dev* 1989; 11: 110–114. [PubMed: 2712233]
10. Stauder JE, Smeets EE, van Mil SG, Curfs LG. The development of visual- and auditory processing in Rett syndrome: an ERP study. *Brain Dev* 2006; 28: 487–494. [PubMed: 16647236]
11. Saunders KJ, McCulloch DL, Kerr AM. Visual function in Rett syndrome. *Dev Med Child Neurol* 1995; 37: 496–504. [PubMed: 7789659]
12. LeBlanc JJ, DeGregorio G, Centofante E et al. Visual evoked potentials detect cortical processing deficits in Rett syndrome. *Ann Neurol* 2015; 78: 775–786. [PubMed: 26332183]
13. Glaze DG. Neurophysiology of Rett syndrome. *J Child Neurol* 2005; 20: 740–746. [PubMed: 16225829]

14. Key AP, Jones D, Peters S. Spoken word processing in Rett syndrome: Evidence from event-related potentials. *Int J Dev Neurosci* 2019; 73: 26–31. [PubMed: 30630072]
15. Peters SU, Katzenstein A, Jones D, Key AP. Distinguishing response to names in Rett and MECP2 Duplication syndrome: An ERP study of auditory social information processing. *Brain Res* 2017; 1675: 71–77. [PubMed: 28870827]
16. Achenbach TM, Ruffle TM. The Child Behavior Checklist and related forms for assessing behavioral/emotional problems and competencies. *Pediatr Rev* 2000; 21: 265–271. [PubMed: 10922023]
17. Squires A, Twombly E, Bricker D, Potter L. *Ages & Stages Questionnaires — 3rd ed.* Baltimore, MD: Paul H. Brookes, 2009.
18. Wilkinson GS, Robertson GJ. *Wide Range Achievement Test—4 (WRAT—4)*. Lutz, FL: Psychological Assessment Resources, 2006.
19. Cuddapah VA, Pillai RB, Shekar KV et al. Methyl-CpG-binding protein 2 (MECP2) mutation type is associated with disease severity in Rett syndrome. *J Med Genet* 2014; 51: 152–158. [PubMed: 24399845]
20. Neul JL, Fang P, Barrish J et al. Specific mutations in methyl-CpG-binding protein 2 confer different severity in Rett syndrome. *Neurology* 2008; 70: 1313–1321. [PubMed: 18337588]
21. Perrin F, Pernier J, Bertrand O, Echallier JF. Spherical splines for scalp potential and current density mapping. *Electroencephalogr Clin Neurophysiol* 1989; 72: 184–187. [PubMed: 2464490]
22. Aso K, Watanabe K, Negoro T et al. Developmental changes of pattern reversal visual evoked potentials. *Brain Dev* 1988; 10: 154–159. [PubMed: 3407851]
23. Cohn NB, Kircher J, Emmerson RY, Dustman RE. Pattern reversal evoked potentials: age, sex and hemispheric asymmetry. *Electroencephalogr Clin Neurophysiol* 1985; 62: 399–405. [PubMed: 2415335]
24. Ehinger Y, Matagne V, Villard L, Roux JC. Rett syndrome from bench to bedside: recent advances. *F1000Res* 2018; 7: 398. [PubMed: 29636907]
25. Durand S, Patrizi A, Quast KB et al. NMDA receptor regulation prevents regression of visual cortical function in the absence of *Mecp2*. *Neuron* 2012; 76: 1078–1090. [PubMed: 23259945]
26. Goffin D, Allen M, Zhang L et al. Rett syndrome mutation *MeCP2* T158A disrupts DNA binding, protein stability and ERP responses. *Nat Neurosci* 2011; 15: 274–283. [PubMed: 22119903]
27. Goffin D, Brodtkin ES, Blendy JA et al. Cellular origins of auditory event-related potential deficits in Rett syndrome. *Nat Neurosci* 2014; 17: 804–806. [PubMed: 24777420]
28. Dong HW, Erickson K, Lee JR et al. Detection of neurophysiological features in female R255X *MeCP2* mutation mice. *Neurobiol Dis* 2020; 145: 105083. [PubMed: 32927061]
29. Ip JPK, Mellios N, Sur M. Rett syndrome: insights into genetic, molecular and circuit mechanisms. *Nat Rev Neurosci* 2018; 19: 368–382. [PubMed: 29740174]
30. Sysoeva OV, Smirnov K, Stroganova TA. Sensory evoked potentials in patients with Rett syndrome through the lens of animal studies: Systematic review. *Clin Neurophysiol* 2020; 131: 213–224. [PubMed: 31812082]
31. Drislane FW. Visual Evoked Potentials. In Blum AS, Rutkove SB (eds): *The Clinical Neurophysiology Primer*. Humana Press 2007.
32. Uren SM, Stewart P, Crosby PA. Subject cooperation and the visual evoked response. *Invest Ophthalmol Vis Sci* 1979; 18: 648–652. [PubMed: 447465]

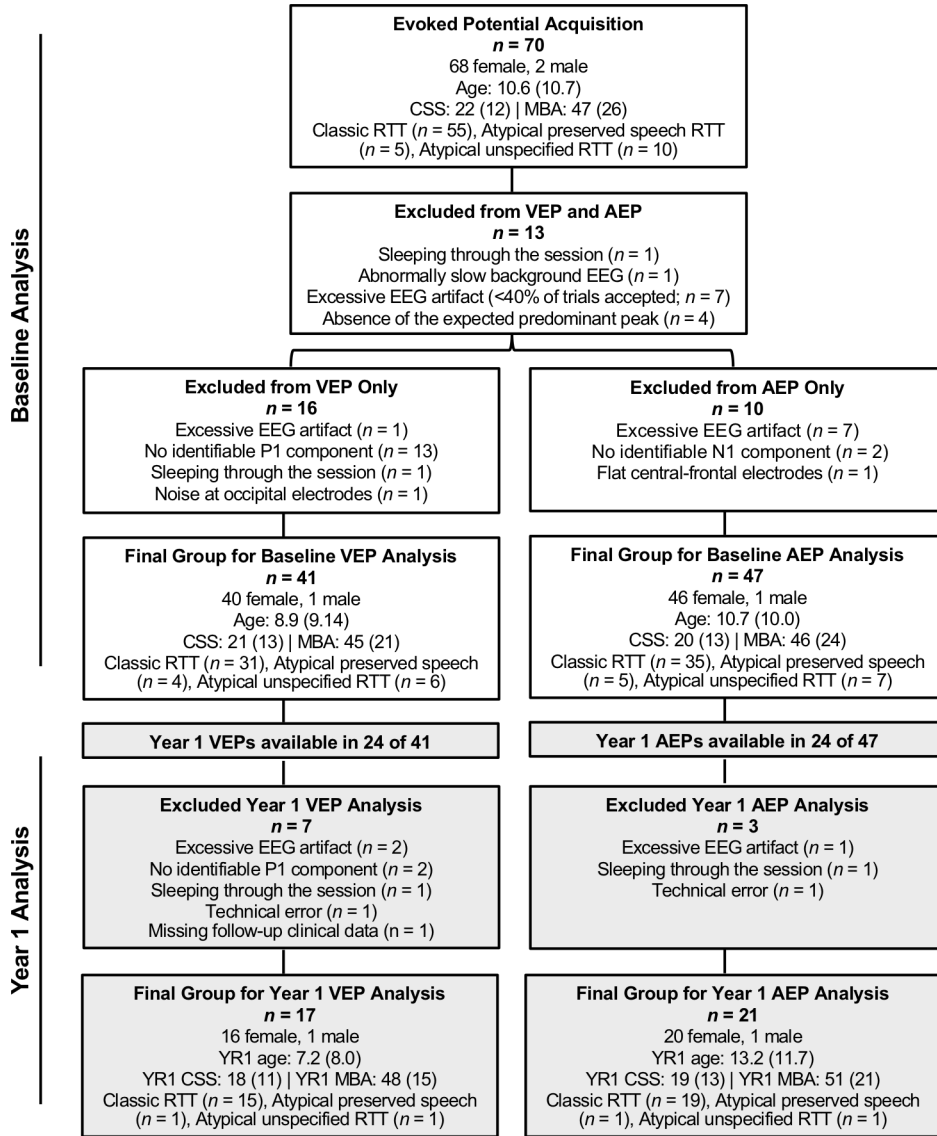
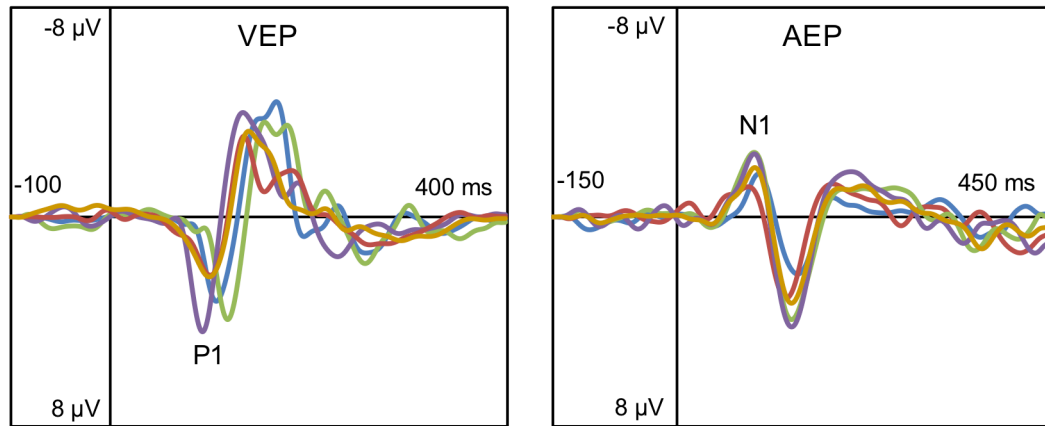
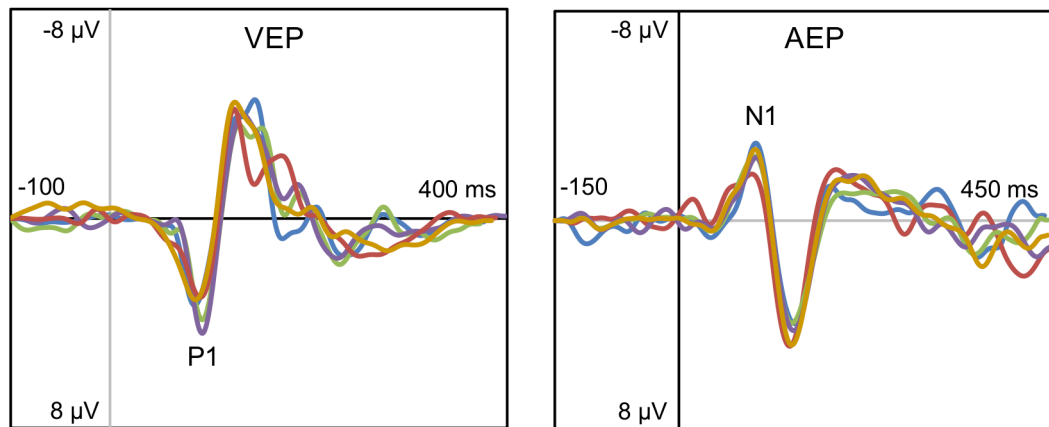


Figure 1. Demographic and clinical information for all participants enrolled in the study, reasons for exclusion, and final samples for the Baseline and Year 1 VEP and AEP analyses. For age, CSS, and MBA, data are presented median (interquartile range). CSS = Clinical Severity Score. MBA = Motor Behavioral Assessment.

a Human Phantom Data No Correction**b** Human Phantom Data with Adjustment

— BCH — CHOP — CCH — VKC — UCD

Figure 2.

(a) Raw VEP and AEP waveforms from the traveling human phantom at each of the five sites. Data from the human phantom was used to determine adjustment factors to better align the latency and amplitude of the VEP/AEP components across sites. The derived adjustment factors were applied to the data for all participants and controls prior to final analysis. (b) VEP and AEP waveforms from the traveling human phantom with the adjustment applied.

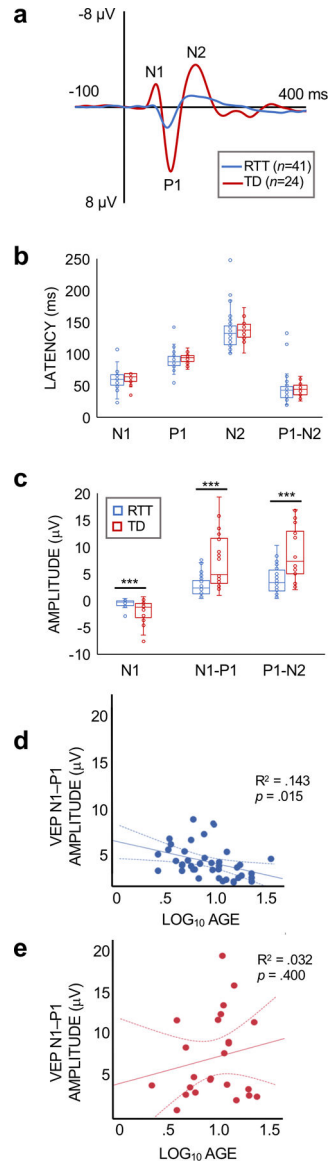


Figure 3.

(a) Grand average VEP waveforms for RTT and TD subjects at electrode Oz. The N1, P1, and N2 components are labeled for the TD waveform. (b) Bar graphs showing the median values and inter-quartile range for the latency of the VEP components for RTT and TD subjects. (c) Bar graphs showing the median values and inter-quartile range for VEP amplitude. The amplitude of the VEP components were significantly lower in the RTT group ($***p < .001$). (d,e) Regression plots for the association for age and VEP amplitude for RTT (d) and TD (e) participants. VEP amplitude declined with age in RTT. No association with age was observed for TD subjects.

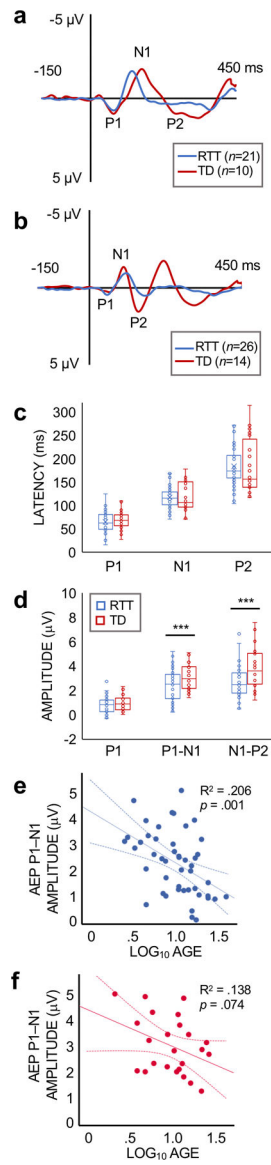


Figure 4.

(a,b) Grand average AEP waveforms for RTT and TD subjects. Due to a strong latency dependence with age, waveforms are plotted separately for younger (a; > 10 years) and older (b; > 10 years;) participants to more accurately display group averages. The P1, N1, and P2 components are labeled for the TD waveform. (c) Bar graph showing the median values and inter-quartile range for the latency of the AEP components for all ages. (d) Bar graph showing the median values and inter-quartile range for the amplitude of the AEP components for all ages. Participants with RTT demonstrated lower P1–N1 and N1–P2 amplitudes compared to TD subjects ($*p < .05$). (e,f) Regression plots for age and AEP amplitude for the RTT (e) and TD (f) groups. AEP P1–N1 amplitude declined with age in participants with RTT. The association between age and P1–N1 amplitude was not significant in the TD group.

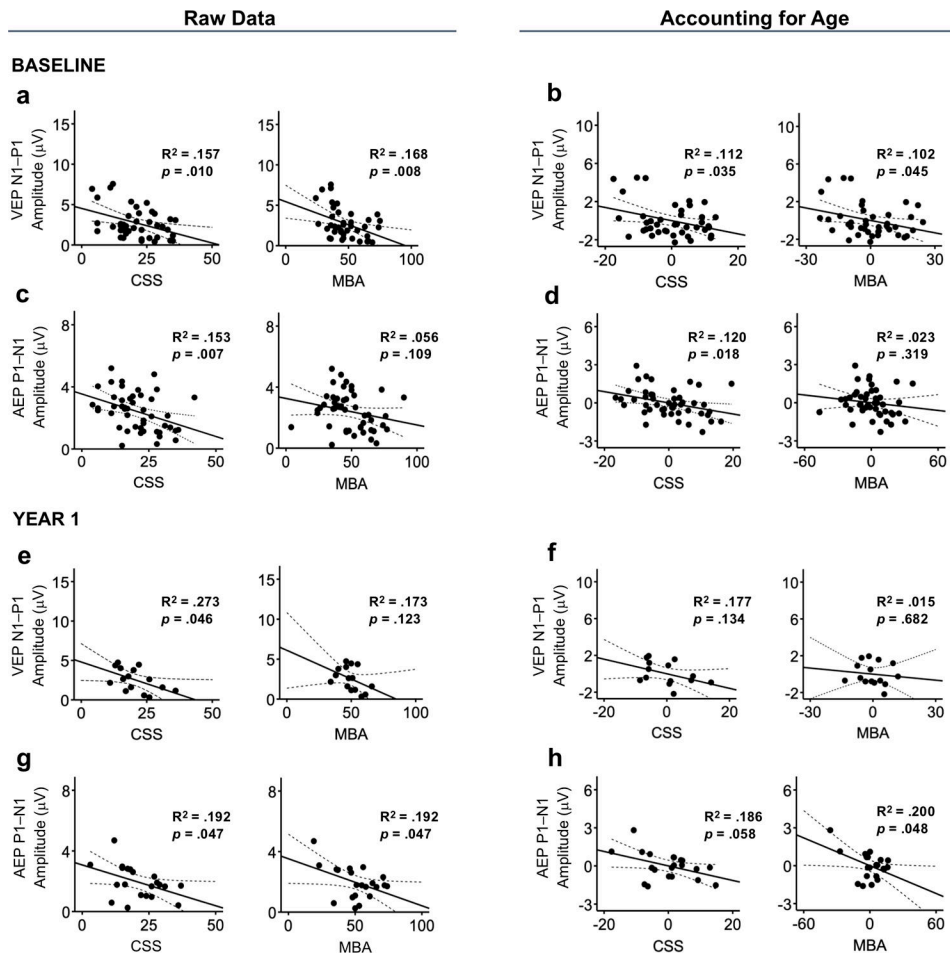
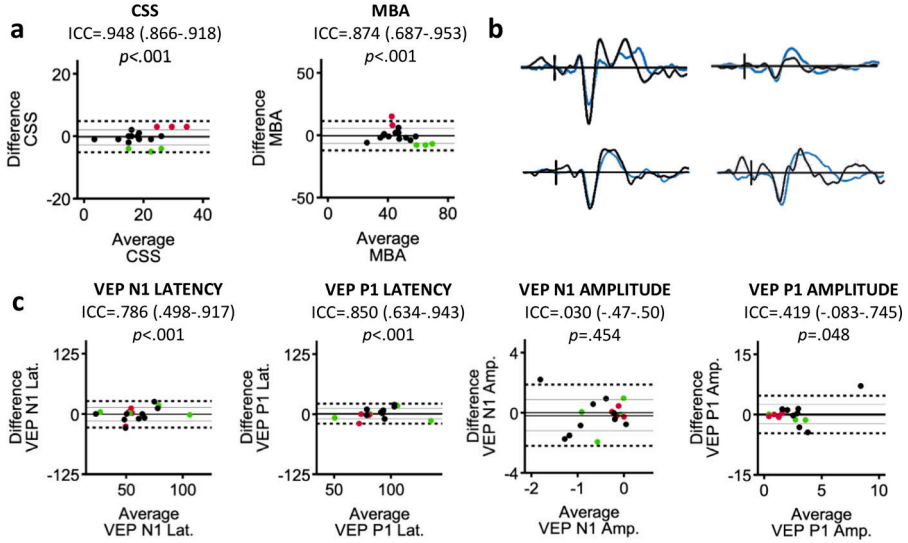


Figure 5. The association between clinical scores (CSS and MBA) and VEP and AEP amplitude for Baseline and Year 1 analyses. Raw associations are shown in the left panel. Partial plots after controlling for the effect of age are shown in the right panel. For the Baseline data, VEP N1–P1 amplitude was negatively associated with both clinical measures (**a**) and this association remained significant after controlling for the effect of age (**b**). AEP P1–N1 amplitude was negatively associated with clinical severity as measured by CSS, but not MBA (**c**, **d**). In the Year 1 data, VEP N1–P1 amplitude was not associated with either clinical measure after accounting for the effect of age (**e**, **f**). AEP P1–N1 amplitude was associated with both measures (**g**), but only the association with CSS remained significant after accounting for age (**h**). Dotted lines represent 95% confidence intervals.

VISUAL EVOKED POTENTIALS



AUDITORY EVOKED POTENTIALS

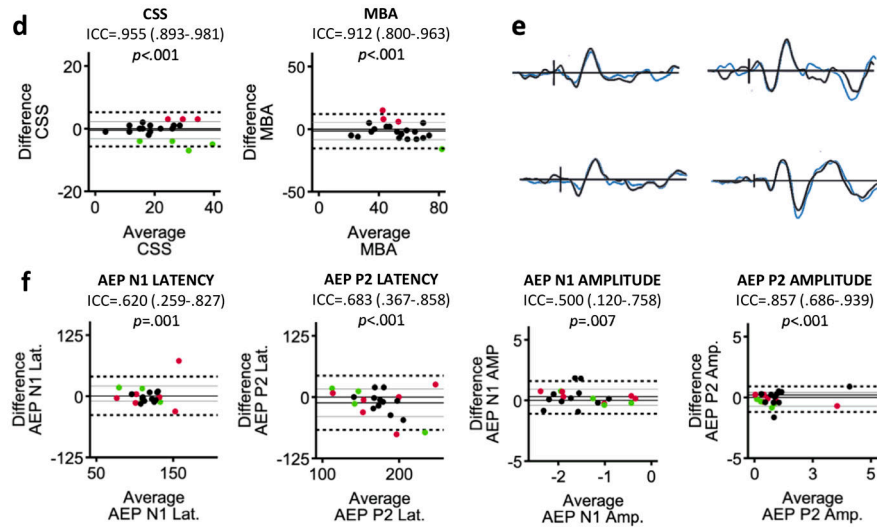


Figure 6.

(a) Bland-Altman plots showing the agreement between clinical severity scores at Baseline and Year 1 for subjects in the analysis of the VEP. Horizontal lines are drawn at the mean difference from Baseline to Year 1 (solid black), the mean difference ± 1 SD (solid gray), and the mean difference ± 1.96 SD (dashed black). Subjects with more than ± 1 SD change in clinical score are highlighted in green (improvement in symptoms) and red (worsening of symptoms). (b) Baseline (black) and Year 1 (blue) waveforms from four individuals with RTT highlighting the repeatability of VEPs acquired one year apart. (c) Bland-Altman plots demonstrating the varying degrees of agreement between Baseline and Year 1 for components of the VEP. (d) Bland-Altman plots showing the agreement in clinical scores for subjects included in the analysis of the AEP. (e) Example waveforms from individual subjects highlighting the repeatability of AEPs. (f) Bland-Altman plots for components of

the AEP. For the VEP and AEP, change in the EPs was not systematically related to changes in clinical severity scores.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1.

Regression coefficients for variables predicting VEP parameters.

	TD	RTT – BASELINE VISIT			RTT – YEAR 1 VISIT*		
	(n = 24)	(n = 41)			(n = 15)		
	Log₁₀ age	Log₁₀ age	CSS	MBA	Log₁₀ age	CSS	MBA
	<i>B</i>	<i>β</i>	<i>β</i>	<i>β</i>	<i>β</i>	<i>β</i>	<i>β</i>
	(<i>p</i>)	(<i>p</i>)	(<i>p</i>)	(<i>p</i>)	(<i>p</i>)	(<i>p</i>)	(<i>p</i>)
N1 Latency	.201 (.347)	.009 (.958)	-.153 (.339)	.012 (.938)	.017 (.951)	.007 (.981)	.333 (.225)
P1 Latency	.049 (.819)	.005 (.976)	-.126 (.433)	.065 (.686)	-.218 (.436)	-.159 (.572)	.073 (.796)
N2 Latency	-.025 (.906)	-.185 (.246)	.028 (.860)	.003 (.985)	-.242 (.384)	-.138 (.623)	.114 (.685)
P1-N2 Time	-.074 (.731)	-.252 (.112)	.123 (.442)	-.040 (.804)	-.191 (.496)	-.061 (.828)	.131 (.642)
N1 Amplitude	-.227 (.286)	.100 (.534)	.357* (.022)	.330* (.035)	.237 (.395)	.444 (.097)	.237 (.395)
N1-P1 Amplitude	.180 (.400)	-.378* (.015)	-.396* (.010)	-.410** (.008)	-.550* (.034)	-.522* (.046)	-.416 (.123)
P1-N2 Amplitude	-.023 (.916)	-.362* (.020)	-.277 (.080)	-.304 (.053)	-.454 (.089)	-.637 (.535)	-.297 (.283)

* Year 1 results are after removing two cases with Cook's D>1 on one or more of the VEP parameters

* $p < .05$ ** $p < .01$

Table 2.

Regression coefficients for variables predicting AEP parameters.

	TD	RTT – BASELINE VISIT			RTT – YEAR 1 VISIT		
	(n = 24)	(n = 47)			(n = 21)		
	Log₁₀ age	Log₁₀ age	CSS	MBA	Log₁₀ age	CSS	MBA
	<i>B</i>	<i>β</i>	<i>β</i>	<i>β</i>	<i>β</i>	<i>β</i>	<i>β</i>
	(<i>p</i>)	(<i>p</i>)	(<i>p</i>)	(<i>p</i>)	(<i>p</i>)	(<i>p</i>)	(<i>p</i>)
P1 Latency	-.757*** (.<.001)	-.364* (.012)	.022 (.885)	.047 (.757)	-.070 (.762)	-.199 (.388)	-.165 (.474)
N1 Latency	-.906*** (.<.001)	-.584*** (.<.001)	-.048 (.749)	-.088 (.555)	-.283 (.214)	-.308 (.174)	-.252 (.270)
P2 Latency	-.807*** (.<.001)	-.523*** (.<.001)	-.063 (.672)	-.070 (.638)	-.420 (.058)	-.467* (.033)	-.304 (.180)
P1 Amplitude	-.393 (.058)	-.429** (.003)	-.417** (.004)	-.311* (.033)	-.078 (.737)	-.314 (.165)	-.456 (.038)
P1–N1 Amplitude	-.371 (.074)	-.454** (.001)	-.392** (.007)	-.237 (.109)	-.109 (.637)	-.438* (.047)	-.438* (.047)
N1–P2 Amplitude	.270 (.203)	-.203 (.172)	-.218 (.141)	-.074 (.619)	.054 (.815)	-.388 (.082)	-.158 (.493)

*
 $p < .05$ **
 $p < .01$ ***
 $p < .001$