Schizophrenia Research 254 (2023) 178–189

Contents lists available at ScienceDirect

Schizophrenia Research

journal homepage: www.elsevier.com/locate/schres

Validation of a suite of ERP and QEEG biomarkers in a pre-competitive, industry-led study in subjects with schizophrenia and healthy volunteers

M. Cecchi [a,*], M. Adachi [b], A. Basile [c], D.L. Buhl [d], H. Chadchankar [e], S. Christensen [f], E. Christian [g], J. Doherty [h], K.C. Fadem [a], B. Farley [h], M.S. Forman [i], S. Honda [b], J. Johannesen [h], B.J. Kinon [j], D. Klamer [k], M.J. Marino [c], C. Missling [k], P. O'Donnell [h], T. Piser [l], C.B. Puryear [m], M.C. Quirk [h], M. Rotte [e], C. Sanchez [n], D.G. Smith [n], J.M. Uslaner [c], D.C. Javitt [o], R.S.E. Keefe [p], D. Mathalon [q], W.Z. Potter [r], D.P. Walling [s], L. Ereshefsky [s,t]

[a] *Cognision, KY, USA*
[b] *Astellas Pharma, CA, USA*
[c] *Merck & Co., Inc., Kenilworth, NJ, USA*
[d] *Takeda Pharmaceuticals, MA, USA*
[e] *Novartis, MA, USA*
[f] *Lundbeck, Denmark*
[g] *Gilgamesh Pharmaceuticals, NY, USA*
[h] *Sage Therapeutics, MA, USA*
[i] *Passage Bio, PA, USA*
[j] *Cyclerion Therapeutics, MA, USA*
[k] *Anavex Life Sciences Corp., NY, USA*
[l] *Onsero Therapeutics, MA, USA*
[m] *Praxis Precision Medicines, MA, USA*
[n] *Alkermes, MA, USA*
[o] *Columbia University, NY, USA*
[p] *VeraSci, NC, USA*
[q] *University of California, San Francisco, CA, USA*
[r] *Independent Consultant, Philadelphia, PA, USA*
[s] *CenExel Research, USA*
[t] *University of Texas Health Science Center at San Antonio, TX, USA*

ARTICLE INFO

Keywords:
ERP Biomarker Qualification Consortium
Event related potentials
Quantitative EEG
Schizophrenia
Test-retest reliability
Automated data analysis

ABSTRACT

Objective: Complexity and lack of standardization have mostly limited the use of event-related potentials (ERPs) and quantitative EEG (QEEG) biomarkers in drug development to small early phase trials. We present results from a clinical study on healthy volunteers (HV) and patients with schizophrenia (SZ) that assessed test-retest, group differences, variance, and correlation with functional assessments for ERP and QEEG measures collected at clinical and commercial trial sites with standardized instrumentation and methods, and analyzed through an automated data analysis pipeline.

Methods: 81 HV and 80 SZ were tested at one of four study sites. Subjects were administered two ERP/EEG testing sessions on separate visits. Sessions included a mismatch negativity paradigm, a 40 Hz auditory steady-state response paradigm, an eyes-closed resting state EEG, and an active auditory oddball paradigm. SZ subjects were also tested on the Brief Assessment of Cognition (BAC), Positive and Negative Syndrome Scale (PANSS), and Virtual Reality Functional Capacity Assessment Tool (VRFCAT).

Results: Standardized ERP/EEG instrumentation and methods ensured few test failures. The automated data analysis pipeline allowed for near real-time analysis with no human intervention. Test-retest reliability was fair-to-excellent for most of the outcome measures. SZ subjects showed significant deficits in ERP and QEEG measures consistent with published academic literature. A subset of ERP and QEEG measures correlated with functional assessments administered to the SZ subjects.

* Corresponding author at: Cognision, 1044 E. Chestnut Street, Louisville, KY 40204, USA.
E-mail address: mcecchi@cognision.com (M. Cecchi).

https://doi.org/10.1016/j.schres.2023.02.018
Received 2 July 2022; Received in revised form 23 November 2022; Accepted 10 February 2023

Conclusions: With standardized instrumentation and methods, complex ERP/EEG testing sessions can be reliably performed at clinical and commercial trial sites to produce high-quality data in near real-time.

1. Introduction

Event-related potentials (ERPs) and quantitative EEG (QEEG) have been established as important translational biomarkers in schizophrenia drug development. When properly implemented, ERPs and QEEG can detect target engagement and response to therapeutic intervention (Javitt et al., 2020; Kantrowitz et al., 2017; Luck et al., 2011; O'Donnell et al., 2013). ERPs in particular, also have shown promise as surrogate endpoints in early-stage procognitive interventional studies (Thomas et al., 2017), and could potentially enable stratification of subjects with schizophrenia by "biotype" (Clementz et al., 2017).

Because of the complexity and lack of standardization for data acquisition and data analysis, the use of ERP and QEEG biomarkers in schizophrenia drug development has been mostly limited to university sites and a few small pharma-sponsored trials. Testing and validating reliable and scalable ERP and QEEG approaches will enable wider use of these measures in drug discovery and development (O'Donnell et al., 2019). An industry-led ERP Biomarker Qualification Consortium (https://erpbiomarkers.org) was constituted with the objective of bringing together industry, academic, and regulatory stakeholders in a spirit of pre-competitive cooperation to ensure that robust and reliable ERP and QEEG biomarkers can be effectively collected in target clinical populations, such as patients with schizophrenia, thus ensuring scalability and consistency across studies. The explicit objectives of the Consortium are to: a) develop and document standardized methods and detailed operating procedures for performing ERP and EEG testing; b) develop a reliable and efficient data analysis pipeline methodology that can be used across studies; c) establish normative ERP and QEEG biomarker metrics in healthy subjects reflective of the population used in Phase 1 safety trials and in a wide range of clinical populations, including schizophrenia; d) quantify and calibrate pharmacodynamic effects on ERP and QEEG biomarkers using well-replicated clinical pharmacological paradigms that mimic the impairment observed in those clinical populations, and e) formally qualify selected ERP and QEEG biomarkers for use in drug trials under the FDA Drug Development Tools Qualification Program. The accomplishment of those objectives will lead to reduced operational risk and trial cost, a more precise estimate of statistical power, and a streamlined regulatory process for trials leveraging qualified ERP and QEEG biomarkers.

This manuscript reports results from the first clinical study sponsored by the Consortium. This was an observational study that recruited healthy volunteers (HV) and subjects with clinically confirmed schizophrenia (SZ). The study a) established mean and variance across cohorts and repeated tests for ERP and QEEG measures collected with a standardized ERP/EEG device; b) validated a predefined, automated data analysis pipeline for ERP and QEEG measures; c) developed normative ERP/QEEG datasets representing SZ subjects and matched HV, and d) quantified the relationship between specific ERP and QEEG parameters and clinically important measures in SZ.

2. Material and methods

2.1. Study design

This was an observational, multicenter study on HV and SZ subjects performed at four study sites in the United States: CenExel-CNS (Torrance, CA), CenExel-CNS (Garden Grove, CA), CenExel-HRI (Marlton, NJ), and the New York State Psychiatric Institute (New York, NY).

The study included 3 visits: Screening, Baseline, and Retest (see Table 1 for an overview of the assessments performed at each visit).

2.2. Study participants

Approximately 20 HV and 20 SZ subjects 21 to 50 years of age were tested at each of the four study sites, for a total of 81 HV and 80 SZ completers. The study (ClinicalTrials.gov number NCT04025502) was approved by institutional review boards for each site. Written informed consent was obtained from each study participant, after which subjects were screened for eligibility.

Inclusion/exclusion criteria selected for the study were consistent with planned schizophrenia trials sponsored by Consortium members.

Eligibility criteria for HV subjects included normal cognitive function as determined by performing within two standard deviations of a normative sample on the Brief Assessment of Cognition (BAC, Atkins et al., 2017; Keefe et al., 2004) Symbol Coding (BAC_SC) and BAC Verbal Memory (BAC_VM). Exclusion criteria included evidence or history of psychiatric illness as determined by The Mini International Neuropsychiatric Interview (MINI, Sheehan et al., 1998), or family history of schizophrenia spectrum disorders in first- or second-degree relatives.

Eligibility criteria for SZ subjects required a diagnosis of schizophrenia as determined by the MINI, excluding all other schizophrenia spectrum disorders. Patients had to be clinically stable, and on a stable regimen of antipsychotic medications for a minimum of 6 weeks, with up to 2 first or second-generation antipsychotics allowed. Exclusion criteria were scores of $\geq$5 for any of the Positive and Negative Syndrome Scale (PANSS, Kay et al., 1987) items P1 (delusions), P3 (hallucinatory behavior), G9 (unusual thought content), and P2 (conceptual disorganization), a score of >6 for the Simpson Angus Scale (SAS, Simpson et al., 1970), or a score of $\geq$6 for the Calgary Depression Scale for Schizophrenia (CDSS, Addington et al., 1990). Eligibility for the PANSS was confirmed at Baseline and Retest.

All subjects were tested for hearing deficits at Screening and had to be able to detect a 1000 Hz tone played at 40 dB in both ears. Subjects also had to pass a saliva drug/alcohol screen at all visits.

Finally, all subjects were required to abstain from medications known to interfere with ERP/EEG assessments within 1 week prior to Screening and throughout the study, and from products containing nicotine and/or caffeine for 60 min prior to ERP/EEG testing.

2.3. ERP and EEG testing sessions

ERP/EEG testing sessions were performed at Baseline and Retest visits and included four tests administered sequentially in rapid succession: 1) a mismatch negativity paradigm, 2) a 40 Hz auditory steady-state response, 3) an eyes-closed resting state EEG, and 4) an active auditory oddball paradigm. Each testing session lasted about 60 min, including headset set up. About halfway through the study, the oddball stimulus in the mismatch negativity paradigm was changed from a frequency-deviant to a duration-deviant. See Table 2 for details of each ERP/EEG test protocol along with pass/fail metrics.

2.4. ERP/EEG data acquisition

ERP and EEG data were recorded using a commercially available, FDA cleared COGNISION® System (Cognision). This system includes all necessary hardware and software to design, perform, and automatically analyze data from ERP/EEG recording sessions. The wireless handheld system is battery-powered and records from active electrodes positioned at Fz, Cz, Pz, F3, P3, F4, and P4 locations of the international 10–20 system (Jasper, 1958). Electrodes were referenced to linked mastoids M1 and M2. Data were digitized at 250 Hz and bandpass filtered from 0.3 to 70 Hz.

Table 1
Study calendar.

Assessments	Visit 1 Screening	Visit 2 Baseline	Visit 3 Retest
Audiometry	X		
EEG/ERP			
Frequency-Deviant MMN[1]		X	X
Duration-Deviant MMN[2]		X	X
Auditory Steady-State response		X	X
Active Auditory Oddball		X	X
Resting EEG		X	X
MINI	X		
BAC[3]	X		
PANSS[4]	X	X[5]	X[5]
CDSS[4]	X		
SAS[4]	X		
VRFCAT[4]	X		

[1] 1st part of study.
[2] 2nd part of study.
[3] Only BAC Symbol Coding & Verbal Memory for HV.
[4] SZ only.
[5] To confirm eligibility.

Stimulus sequences were controlled using the COGNISION® Software. Auditory stimuli were generated within the COGNISION® Headset and presented binaurally through integrated medical-grade insert earphones. Task responses were captured using the integrated response buttons on the COGNISION® Handset. Additional information about the COGNISION® System is described in Cecchi et al., 2015.

2.5. Data quality review

At the end of each testing session, data were immediately available for quality review through the web-enabled COGNISION® System software. Quality technicians blind to demographics and diagnostic information evaluated the data against predefined objective quality metrics, and those tests which passed quality review were flagged for automatic analysis. Details of the quality metrics for each ERP/EEG test protocol are shown in Table 2.

2.6. Automated data cleaning, preprocessing, and feature extraction

For tests that passed QC, data cleaning, preprocessing, and extraction of ERP/QEEG parameters were automatically performed with the COGNISION® Software through a predefined data and statistical analysis pipeline (see Fig. 1).

Consistent with procedures commonly employed in pharma-sponsored clinical trials, ERP/QEEG endpoints were predefined in the study protocol and statistical analysis plan. Artifact detection, preprocessing, and feature definitions for the ERP/QEEG endpoints are shown in Table 3.

2.7. Functional assessments

Functional assessments were administered by trained test administrators using the VeraSci electronic Clinical Outcomes Assessment (eCOA) Pathway platform (now part of the WCG eCOA/ePRO platform). See Table 1 for a list of the functional assessments.

2.8. Statistical analysis

Group differences between HV and SZ subjects for demographic characteristics and functional assessments were analyzed using a two-tailed *t*-test for continuous variables (age, education, BAC_SC and BAC_VM), and a chi-squared test for categorical variables (race and sex).

Group differences between HV and SZ subjects for ERP/QEEG endpoints were analyzed using a repeated measures ANOVA with group and visit as factors. Effect size was calculated as Cohen's *d*. When a significant group-by-visit interaction was found, a Bonferroni post hoc analysis was performed to test for possible group differences at each visit.

Baseline/retest variability for ERP/QEEG endpoints was calculated separately for HV and SZ groups as Intraclass Correlation Coefficients (ICC).

Finally, correlations between ERP/QEEG endpoints and functional assessments in SZ were analyzed using Pearson correlation coefficients. To limit the number of comparisons, only ERP/QEEG endpoints that showed HV vs SZ differences were included. Furthermore, Inter-Trial Coherence (ITC) correlation analyses were restricted to the 301–400

Table 2
EEG/ERP test descriptions and quality metrics.

EEG/ERP Test	Paradigm & Stimulus[1]	Sequence	Quality Review[2]	Total/ Passed
FD-MMN	Auditory Oddball Standard = 1000 Hz, 100 ms, 90 %, 85 dB Deviant = 2000 Hz, 100 ms, 10 %, 85 dB	Stimuli presented in pseudorandom order so that 6 to 12 standards were presented between deviants for a total of 1200 stimuli. The interstimulus interval was 600 ms.	Ensure that data from all 7 channels are present, 1200 epochs were collected, recording is not contaminated with 60 Hz powerline noise, and that a clear ERP can be distinguished.	176/ 176
DD-MMN	Auditory Oddball Standard = 1000 Hz, 50 ms, 90 %, 85 dB Deviant = 1000 Hz, 100 ms, 10 %, 85 dB	Stimuli presented in pseudorandom order so that 6 to 12 standards were presented between deviants for a total of 2000 stimuli. The interstimulus interval was 600 ms.	Ensure that data from all 7 channels are present, 2000 epochs were collected, recording is not contaminated with 60 Hz powerline noise, and that a clear ERP can be distinguished.	146/ 145
ASSR	Stimulus Train 500 ms duration, 40 Hz white noise click trains, 85 dB	Click trains presented every 1000 ms for a total of 200 repetitions.	Ensure that data from all 7 channels are present, 200 epochs were collected, recording is not contaminated with 60 Hz powerline noise, and that a clear ERP can be distinguished.	322/ 311
Active Oddball	Auditory Oddball Standard = 1000 Hz, 100 ms, 80 %, 85 dB Deviant = 2000 Hz, 100 ms, 20 %, 85 dB	Stimuli presented in pseudorandom order so that 2 to 5 standards were presented between deviants for a total of 300 stimuli. The interstimulus interval was randomized between 2500 and 3000 ms. Subjects were instructed to press a button on the ERP/EEG recording device as soon as possible each time they heard the deviant (target) stimulus.	Ensure that data from all 7 channels are present, 300 epochs were collected, recording is not contaminated with 60 Hz powerline noise, a clear ERP can be distinguished, and that there are at least 20 epochs with correct button presses.	322/ 299
Resting EEG	Eyes Closed n/a	Subjects were instructed to rest with their eyes closed for 5 min of EEG recording.	Ensure that data from all 7 channels are present, at least 180 s are recorded, and that recording is not contaminated with 60 Hz powerline noise.	322/ 321

Abbreviations: FD-MMN = Frequency Deviant Mismatch Negativity; DD-MMN = Duration Deviant Mismatch Negativity; ASSR = Auditory Steady-State Response.
[1] All auditory stimuli were presented binaurally through medical grade insert earphones.
[2] All tests were reviewed by a quality technician blind to demographics and diagnostic information to determine if the test "passed" the quality review described in the table.

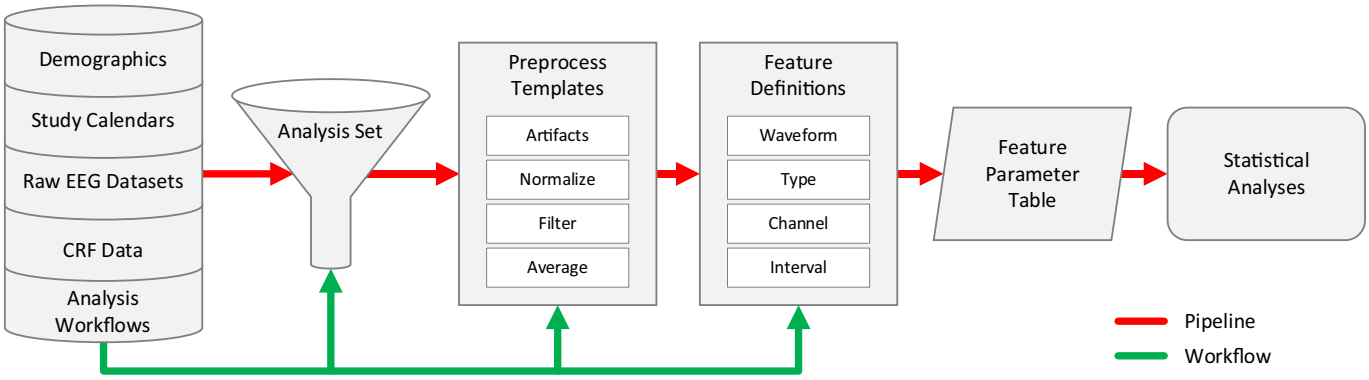

Fig. 1. All study data including subject info, calendar and visit info, raw EEG test data, and data from the study case report forms (ICH Harmonised Tripartite Guideline, 2016), along with all analysis workflows are managed in a secure web-based SQL database. Clinical and psychometric data can also be captured in the database and used in subsequent analyses. The analysis workflows include predefined preprocess templates which determine how each EEG/ERP dataset is preprocessed along with feature definition algorithms which define exactly how each feature will be extracted from the preprocessed EEG/ERP data (see Table 3). The statistical analyses are also predefined to act on the automatically generated feature parameter tables. To generate the feature parameter and statistical analysis output, the user tags the appropriate subjects as part of an "analysis set" (which has been predefined in the study protocol) and initiates the analysis on those subjects to generate the output. The saved analysis workflows can be reused across studies to facilitate standardization and comparisons between different study cohorts.

ms latency block, as it was the latency interval with the largest SZ deficits. Parameters from functional assessments included in the correlation analyses comprised of Digit Sequencing (BAC_DS), BAC_SC, Tower of London (BAC_TL), Token Motor (BAC_TM), Verbal Fluency Total (BAC_VF) and BAC_VM from the BAC, Total Score (PANSS_TOT), Positive Symptoms Subscale (PANSS_PS), and Negative Symptoms Subscale (PANSS_NS) from the PANSS, and Adjusted Total Time (VRFCAT_AT), Total Forced Progressions (VRFCAT_FP), and Total Error Count (VRFCAT_TE) from the VRFCAT (Keefe et al., 2016).

3. Results

3.1. Demographics and clinical characteristics

There were no significant differences in age, gender, and race between HV and SZ groups. As expected, SZ subjects had significantly fewer years of education ($t = 5.271$, $p < 0.01$). In addition, SZ subjects had significantly lower scores than HV in BAC_SC ($t = 4.848$, $p < 0.01$) and BAC_VM ($t = 5.657$, $p < 0.01$) (see Table 4).

3.2. Data quality

Data quality review resulted in the following number of tests that passed quality control criteria: 176 out of 176 for frequency-deviant mismatch negativity; 145 out of 146 for duration-deviant mismatch negativity; 311 out of 322 for auditory steady-state response; 299 out of 322 for auditory active oddball, and 321 out of 322 for resting EEG (see Table 2). The most common reason for test rejection was the inability of SZ subjects to perform the task associated with the active oddball as instructed.

3.3. ERP and QEEG comparisons in HV vs SZ subjects

Statistical significance, effect size, and test-retest reliability for prespecified ERP/QEEG endpoints are summarized in Table 5.

A breakdown of all ICC values for ERP/QEEG measures by study site is shown in supplementary Table S1.

ERP/QEEG endpoints from individual tests, as well as additional ERP parameters that may be of interest to the community but were not part of the prespecified study endpoints for the study are made available by the ERP Biomarker Qualification Consortium at Shared Data Folder.

3.4. Frequency-deviant mismatch negativity

Grand average waves for standard and deviant stimuli and the grand difference wave (deviant-standard) for the passive, frequency-deviant auditory oddball in HV vs SZ subjects are shown in Fig. 2.

Statistical comparisons for the N100 from the standard stimulus showed a lower peak amplitude at Baseline than at Retest ($F_{1,86} = 4.86$, $p < 0.05$) with no group effects or significant interactions.

Statistical analyses of ERP parameters from the difference wave revealed a reduced P3a amplitude ($F_{1,86} = 8.531$, $p < 0.01$), and a prolonged P3a latency ($F_{1,86} = 4.474$, $p < 0.05$) in SZ subjects compared to HV.

3.5. Duration-deviant mismatch negativity

Grand average waves for standard and deviant stimuli and the grand difference wave (deviant-standard) for the passive, duration-deviant auditory oddball in HV vs SZ subjects are shown in Fig. 3.

Statistical comparisons for the N100 from the standard stimulus showed a significant group-by-visit interaction for the peak amplitude ($F_{1,70} = 5.061$, $p < 0.05$) with no significant differences observed in the subsequent post hoc analysis.

Statistical analyses of ERP parameters from the difference wave revealed significantly reduced MMN amplitude ($F_{1,70} = 5.573$, $p < 0.05$) and P3a amplitude ($F_{1,70} = 8.529$, $p < 0.01$) in SZ subjects compared to HV.

3.6. Auditory steady-state response

Grand average waves generated during the ASSR paradigm, along with frequency-vs-time Inter-Trial Coherence (ITC) and Evoked Power (EP) plots in HV and SZ subjects are shown in Fig. 4.

Statistical analyses revealed lower ITC for SZ in the 1-500 ms stimulus window ($F_{1,145} = 8.674$, $p < 0.01$). When data was analyzed for 100 ms latency intervals (Light et al., 2006), SZ had lower ITC at intervals 101-200 ms ($F_{1,145} = 4.057$, $p < 0.05$), 201–300 ($F_{1,145} = 5.171$, $p < 0.05$), 301–400 ($F_{1,145} = 6.497$, $p < 0.05$), and 401–500 ($F_{1,145} = 4.461$, $p < 0.05$). There was also a trend toward significance for the 1-100 ms interval ($F_{1,145} = 3.738$, $p = 0.055$). As would be expected, no significant group differences were found for ITC at the prestimulus interval (−99–0 ms).

Finally, EP showed a mild decrease in SZ that was however not statistically significant ($F_{1,145} = 1.859$, $p = 0.175$).

3.7. Active auditory oddball

Grand average waves for standard and deviant stimuli from the active auditory oddball in HV vs SZ subjects are shown in Fig. 5.

Statistical comparisons for ERP parameters showed significant group ($F_{1,140} = 13.148$, $p < 0.01$) and visit ($F_{1,140} = 8.945$, $p < 0.01$) effects for

Table 3
ERP and QEEG artifact detection, preprocessing, and feature parameter definitions.

EEG/ERP Test	Artifact definitions[1]	Feature	Type[2]	Filter[3]	Stim[4]	Chan	Interval 1[5]	Interval 2[6]
FD-MMN	EEG voltages > ±100 μV or >1.8 × VRMS.	N100	Peak	50	1	Cz	56–153	52–164
	A mean of 243 out of 1200 total epochs were detected as artifacts in each "valid" data-set.	MMN	Peak	30	2–1	Fz	92–240	88–244
		P3a	Peak	30	2–1	Cz	196–352	192–356
DD-MMN	EEG voltages > ±100 μV or >1.8 × VRMS.	N100	Peak	50	1	Cz	64–132	60–136
	A mean of 608 out of 2000 total epochs were detected as artifacts in each "valid" data-set.	MMN	Peak	30	2–1	Fz	104–260	100–264
		P3a	Peak	30	2–1	Cz	224–352	220–356
ASSR	EEG voltages > ±200 μV or >2 × VRMS.	ITC000	ITC	38–42	1	Fz	−99–0	
	A mean of 22 out of 200 total epochs were detected as artifacts in each "valid" data-set.	ITC100	ITC	38–42	1	Fz	1–100	
		ITC200	ITC	38–42	1	Fz	101–200	
		ITC300	ITC	38–42	1	Fz	201–300	
		ITC400	ITC	38–42	1	Fz	301–400	
		ITC500	ITC	38–42	1	Fz	401–500	
		ITC1500	ITC	38–42	1	Fz	1–500	
		EP	EP	38–42	1	Fz	1–500	
Active Oddball	EEG voltages > ±150 μV or >2 × VRMS. Also missed button presses to the target tone or false alarm presses to the standard tone.	N100	Peak	50	1	Cz	68–140	64–144
		P3b	Peak	16	2	Pz	244–472	244–476
	A mean of 68 out of 300 total epochs were detected as artifacts in each "valid" data-set.	BPA	TA		2			
		MRT	MRT		2			
Resting EEG[7]	Pharmaco-EEG analysis was performed using the Welch method (Barbé et al., 2010) with 4 s windows and 50 % overlap. Windows with EEG voltages > ±100 μV were detected as artifacts in each "valid" data-set.	Delta-Abs	AP			AVG	[1.5–6]	
		Delta-Rel	RP			AVG	[1.5–6]	
		Theta-Abs	AP			AVG	[6–8.5]	
		Theta-Rel	RP			AVG	[6–8.5]	
		Alpha1-Abs	AP			AVG	[8.5–10.5]	
		Alpha1-Rel	RP			AVG	[8.5–10.5]	
		Alpha2-Abs	AP			AVG	[10.5–12.5]	
		Alpha2-Rel	RP			AVG	[10.5–12.5]	
		Beta1-Abs	RP			AVG	[12.5–18.5]	
		Beta1-Rel	RP			AVG	[12.5–18.5]	
		Beta2-Abs	AP			AVG	[18.5–21]	
		Beta2-Rel	RP			AVG	[18.5–21]	
		Beta3-Abs	AP			AVG	[21−30]	
		Beta3-Rel	RP			AVG	[21–30]	
		TotalPow	AP			AVG	[1.5–30]	
		Gamma-Abs	AP			AVG	[30–40]	
		AlphaPeak	PAF			AVG	[6–12.5]	
		SlowWave	SWI			AVG		
		ThetaBeta	TBR			AVG		

Abbreviations: FD-MMN = Frequency Deviant Mismatch Negativity; DD-MMN = Duration Deviant Mismatch Negativity; ASSR = Auditory Steady-State Response; ITC = Intertiral Coherence; EP = Evoked Power; TA = Task Accuracy; MRT = Median Reaction Time; PAF = Peak Alpha Frequency; SWI = Slow Wave Index; TBR = Theta/Beta Ratio; RP = Relative Power; AP = Absolute Power.

[1] Artifact detection and removal are performed fully automatically as part of the data analysis pipeline. Any epochs with artifacts occurring on any channel were removed from all feature analyses.

[2] Peak = a maximum or minimum within a time window; Inter-Trial Coherence = an estimation of the strength of phase locking of the EEG signals across individual trials independent of the signal amplitude (Light et al., 2006); Evoked Power = the power in the 40 Hz band computed from the average ERP; Task Accuracy = the ratio of the total number of correct button presses to the target stimuli to the total number of target stimuli; Median Reaction Time = the median time from target stimulus onset to detection of correct button press; Absolute Power = the total power in a frequency band; Relative Power = the ratio of the power in a frequency band to the total power for all bands; Frequency Peak = frequency with maximum power in a frequency range; Slow-wave Index = the ratio between Alpha activity and the sum of the activity in the Delta and Theta frequency bands; Theta/Beta Ratio = the ratio between Theta activity and the activity in the Beta1 and Beta2 frequency bands.

[3] Lowpass for all ERP paradigms except for ASSR which is a passband.

[4] Defines from which stimulus the ERP is created. Stim 1 is the "standard" for all oddball paradigms. Stim 2 is the "deviant" for all oddball paradigms.

[5] For oddball paradigms. This is an initial window interval in ms for detecting a peak. For the ASSR paradigm, it is the time window in ms where Intertrial Coherence (ITC) and Evoked Power (EP) are calculated. For the Pharmaco-EEG features, it is the passband in Hz.

[6] A secondary window interval if a peak was not detected in "Interval 1".

[7] The frequency bands selected for Pharmaco-EEG analysis were taken from an International Pharmaco-EEG Society (IPEG) guidance document intended to standardize Pharmaco-EEG analysis methods for the pharmaceutical industry (Jobert et al., 2012).

P3b amplitude. The peak amplitude was lower in SZ subjects than HV, and at Retest when compared to Baseline.

When performance in the behavioral response was analyzed, SZ subjects showed lower button press accuracy ($F_{1,142} = 12.030$, $p < 0.01$) and delayed median reaction time ($F_{1,142} = 21.052$, $p < 0.01$) compared to HV.

3.8. Eyes-closed resting EEG

Power spectral densities for HV and SZ subjects are shown in Fig. 6.

Statistical analyses for the QEEG parameters showed higher absolute Delta power ($F_{1,158} = 8.527$, $p < 0.01$), lower relative Beta1 ($F_{1,158} = 6.581$, $p < 0.05$) and Beta2 ($F_{1,158} = 5.691$, $p < 0.05$) power, and higher Theta/Beta ratio ($F_{1,158} = 7.065$, $p < 0.01$) in SZ subjects. There was also a significant group × visit interaction for Theta relative power that, after subsequent post hoc analysis, revealed higher Theta relative power in SZ at Retest.

3.9. Correlation of parameters from ERP and QEEG testing with functional assessments in subjects with schizophrenia

Table 6 summarizes significant correlations between ERP and QEEG parameters that showed changes in SZ subjects, and functional assessments.

For the passive, frequency-deviant auditory oddball, P3a amplitude correlated with PANSS_PS, P3a latency correlated with BAC_SC and VRFCAT_FP. For the passive, duration-deviant auditory oddball, MMN and P3a amplitude correlated with BAC_SC. For the auditory steady-state response, ITC in the 301-400 ms latency window correlated with BAC_DS and BAC_VF. For the active oddball median reaction time

Table 4
Demographics and clinical characteristics.

	Healthy volunteers (HV)	Patients (SZ)
Sample size	81	80
Age[1]	37.27 (1.08)	38.40 (0.87)
Gender		
Male[2]	45 (55.6 %)	49 (61.3 %)
Female[2]	36 (44.4 %)	31 (38.8 %)
Race		
White[2,3]	14 (17.3 %)	9 (11.3 %)
African American[2,3]	31 (38.3 %)	42 (52.5 %)
Other Race[2,3]	36 (44.4 %)	29 (36.3 %)
Education[1]	13.90 (0.23)	12.21 (0.17)**
Duration of Illness[1,4]	n/a	14.54 (0.85)
CDSS[1,5]	n/a	1.24 (0.18)
SAS[1,5]	n/a	0.275 (0.085)
BAC Verbal Memory[1]	43.36 (0.85)	35.26 (1.16)**
BAC Symbol Coding[1]	48.67 (1.10)	40.03 (1.33)**
BAC Digit Sequencing[1,4]	n/a	15.48 (0.49)
BAC Token Motor[1,4]	n/a	66.70 (3.23)
BAC Tower of London[1,4]	n/a	40.99 (1.18)
BAC Verbal Fluency[1,4]	n/a	12.56 (1.40)
PANSS Total Score[1,4]	n/a	61.38 (1.35)
PANSS Positive Symptoms[1,4]	n/a	14.89 (0.50)
PANSS Negative Symptoms[1,4]	n/a	16.41 (0.45)
VRFCAT Adjusted Total Time (s)[1,4]	n/a	761.45 (21.08)
VRFCAT Total Error Count[1,4]	n/a	2.33 (0.33)
VRFCAT Total Forced Progressions[1,4]	n/a	46.50 (1.83)

Abbreviations: CDSS = Calgary Depression Scale for Schizophrenia; SAS = Simpson-Angus Scale; BAC = Brief Assessment of Cognision; PANSS = Positive and Negative Syndrome Scale; VRFCAT = Virtual Reality Functional Capacity Assessment Tool.

[1] Mean (±SEM).

[2] Total (% of Total).

[3] Racial labels from FDA, 2016, Guidance for Industry, Collection of Race and Ethnicity Data in Clinical Trials (FDA, 2016).

[4] SZ only.

[5] A breakdown of participants' demographics by study site is shown in supplementary Table S2.

** $p < 0.01$

correlated with PANSS_NS, and button press accuracy with VRFCAT_AT. Finally, for the resting EEG, Delta absolute power correlated with PANSS_PS, while Theta/Beta ratio correlated with BAC_VF and VRFCAT_FP.

4. Discussion

The current study provides evidence that, with standardized equipment and methods, complex ERP/EEG testing sessions can be reliably performed across clinical and commercial trial sites to produce high-quality data with few test failures. In our study, group differences reproduced results reported in the academic literature, and test-retest reliability was fair-to-excellent for most of the measures collected. Also, several ERP and QEEG measures exhibited significant correlations with functional measures.

4.1. Automated pipelined data analysis and data quality

ERP/QEEG parameters were obtained through a predefined, automated data analysis pipeline. This analysis approach contributed to the high quality of the data by ensuring that data cleaning, preprocessing, and feature extraction were consistent across datasets and free from subjective interpretation. The analysis pipeline output is available immediately at the end of each testing session. For interventional trials, this rapid data throughput and analyses will enable quality and training interventions, as well as facilitate study protocol changes in adaptive trial designs.

Because early phase clinical trials of experimental compounds are performed on small cohorts, it is important that the ERP/EEG testing sessions be performed with very few test failures. In the current study, the number of ERP and EEG tests that did not meet quality criteria was very low, even when compared to similar studies performed at specialized academic ERP labs (see for example Light et al., 2014, and Turetsky et al., 2007). The most common reason for test rejection was the inability of a subset of SZ subjects to correctly perform the task associated with the active oddball test, a factor that should be taken into account when designing SZ clinical trials that includes active ERP paradigms.

An additional requirement for accurate assessments of the effects of experimental drugs is adequate test-retest reliability. A large number of ERP/QEEG parameters analyzed in this study had a test-retest reliability score that was good or excellent (Cicchetti and Sparrow, 1981), with similar reliability patterns across sites. Reliability studies from academic labs have occasionally reported higher ICC values for some ERP measures (Light and Braff, 2005; Turetsky et al., 2007). Those studies were performed at specialized ERP labs under conditions that could not be controlled at commercial trial sites. Our study implementation mimics multicenter pharma trials, where sites might not have dedicated spaces explicitly set up for ERP recordings, and test administrators are often not trained electrophysiologists and rotate throughout the study so that a subject may have different test administrators across sessions. A more relevant comparison for our results is with multicenter academic studies like the NAPLS study (Roach et al., 2020). That study reported test-retest reliability values for the MMN peak that are similar to ours for amplitude, and slightly less compelling for latency.

4.2. ERP and QEEG comparisons in HV vs SZ subjects

Our findings closely match published results from top academic labs.

SZ subjects showed a reduction in MMN amplitude from the duration-deviant paradigm but not from the frequency-deviant. Patient functional status is an important determinant of the pattern of MMN dysfunction in SZ, such that deficits in duration MMN appear to be present across all SZ subjects, whereas deficits in frequency MMN are restricted to a subgroup of low-functioning subjects drawn primarily from supervised residential-care settings (see for example Lee et al., 2017), and may index reductions in auditory cortex volume that are observed during initial years of the disease (Salisbury et al., 2007). The bimodal distribution of tone matching impairments in SZ suggests that these may represent an etiologically distinct subgroup (Dondé et al., 2019). Our results add to the evidence that relatively high functioning patients drawn from outpatient settings show deficits in duration MMN but relatively intact frequency discrimination. The differential MMN findings may also reflect the differential structural correlations of the different MMN types (Curtis et al., 2021).

While group differences in MMN amplitude were dependent on the kind of deviant, P3a amplitude was decreased in SZ subjects for both frequency and duration-deviant paradigms. Deficits in P3a amplitude have been a consistent finding in SZ, and are present even at the early stages of the disease (Ford et al., 2010; Light et al., 2014; Mathalon et al., 2000; Nagai et al., 2013). Interestingly, SZ subjects also showed a significant increase in P3a latency for the frequency-deviant. Though similar data have been previously reported (Frodl et al., 2001; Li et al., 2013), this is a relatively novel finding, and its specificity to the frequency-deviant paradigm further suggests differential patterns of dysfunction according to deviant type.

ASSR test results showed a significant decrease in ITC in SZ subjects for the 1–500 ms stimulus window, and for most of the 100 ms latency blocks collected during stimulus presentation. Similar to Light et al., 2006, the largest group difference was observed for the 301–400 ms latency block.

ASSR EP was not significantly different between groups. ASSR ITC and EP do not necessarily covary (Roach and Mathalon, 2008). Though most studies have shown a decrease in both ITC and EP in SZ subjects (Thuné et al., 2016), a decrease in ITC in the absence of EP deficits has also been reported (Hirano et al., 2015; Kirihara et al., 2012). The reason for this discrepancy is unclear, but it is likely not the result of methodological differences across studies, as stimulus characteristics and

Table 5
EEG/ERP features, significance, effect sizes, and interclass correlation coefficients.

EEG/ERP Test	Feature	Measure (units)	HV_{AVG}	HV_{BL}	HV_{RT}	SZ_{AVG}[1]	SZ_{BL}	SZ_{RT}	Cohen's D	ICC HV[2]	ICC SZ[2]
FD-MMN	N100-A	Amplitude (μV)	−1.19 (0.14)	−1.03 (0.18)	−1.36 (0.21)	−1.21 (0.16)	−1.11 (0.22)	−1.31 (0.25)	0.014	0.582†	0.776†††
	N100-L	Latency (ms)	110 (1.5)	110 (1.8)	110 (2.3)	110 (1.8)	114 (2.1)	107 (2.7)	0.030	0.600††	0.454†
	MMN-A	Amplitude (μV)	−4.37 (0.22)	−4.38 (0.31)	−4.35 (0.32)	−4.5 (0.20)	−4.47 (0.30)	−4.53 (0.29)	0.065	0.552†	0.630††
	MMN-L	Latency (ms)	146 (2.7)	146 (3.9)	147 (3.8)	153 (2.4)	150 (3.9)	157 (2.8)	0.292	0.460†	0.334
	P3a-A	Amplitude (μV)	3.26 (0.19)	3.27 (0.29)	3.24 (0.26)	2.48 (0.15)**	2.50 (0.20)	2.45 (0.21)	0.491	0.303	0.092
	P3a-L	Latency (ms)	261 (4.2)	264 (5.7)	257 (6.1)	275 (4.3)*	277 (6.1)	273 (6.1)	0.354	0.228	0.228
DD-MMN	N100-A	Amplitude (μV)	−1.21 (0.19)	−1.36 (0.29)	−1.05 (0.29)	−1.09 (0.16)	−1.02 (0.23)	−1.16 (0.23)	0.075	0.669††	0.864†††
	N100-L	Latency (ms)	87 (1.4)	85 (2.0)	88 (1.9)	89 (1.9)	90 (2.7)	88 (2.7)	0.142	0.315	0.707††
	MMN-A	Amplitude (μV)	−5.58 (0.31)	−5.79 (0.46)	−5.50 (0.41)	−4.44 (0.27)*	−4.29 (0.37)	−4.59 (0.32)	0.475	0.702††	0.556†
	MMN-L	Latency (ms)	173 (3.1)	174 (4.8)	171 (4.0)	164 (3.1)	164 (5.1)	163 (3.7)	0.343	0.784†††	0.639††
	P3a-A	Amplitude (μV)	6.72 (0.47)	6.99 (0.68)	6.57 (0.67)	4.52 (0.33)**	4.61 (0.45)	4.44 (0.49)	0.635	0.762†††	0.826†††
	P3a-L	Latency (ms)	278 (3.2)	279 (4.6)	277 (4.7)	277 (3.6)	274 (4.2)	279 (5.8)	0.045	0.669††	0.554†
ASSR	ITC000	ITC	0.073 (0.002)	0.075 (0.003)	0.071 (0.003)	0.068 (0.002)	0.069 (0.002)	0.067 (0.003)	0.185	−0.053	0.053
	ITC100	ITC	0.176 (0.007)	0.175 (0.011)	0.178 (0.010)	0.150 (0.005)	0.147 (0.007)	0.152 (0.007)	0.342	0.475†	0.507†
	ITC200	ITC	0.341 (0.013)	0.332 (0.019)	0.350 (0.016)	0.290 (0.011)*	0.290 (0.016)	0.289 (0.015)	0.352	0.658††	0.804†††
	ITC300	ITC	0.403 (0.014)	0.396 (0.020)	0.410 (0.018)	0.337 (0.013)*	0.340 (0.019)	0.334 (0.018)	0.401	0.710††	0.767†††
	ITC400	ITC	0.363 (0.013)	0.361 (0.019)	0.365 (0.018)	0.294 (0.012)*	0.302 (0.017)	0.287 (0.017)	0.446	0.689††	0.753†††
	ITC500	ITC	0.350 (0.012)	0.345 (0.018)	0.354 (0.017)	0.292 (0.012)*	0.300 (0.018)	0.284 (0.017)	0.379	0.612††	0.720††
	ITC1500	ITC	0.433 (0.012)	0.425 (0.018)	0.441 (0.017)	0.360 (0.012)**	0.366 (0.018)	0.354 (0.016)	0.486	0.675††	0.799†††
	EP	Power (μV^2/Hz)	0.198 (0.016)	0.198 (0.023)	0.197 (0.023)	0.152 (0.010)	0.163 (0.016)	0.142 (0.012)	0.273	0.604††	0.518†
Active Oddball	P3b-A	Amplitude (μV)	9.57 (0.34)	9.97 (0.48)	9.31 (0.50)	7.39 (0.32)**	8.10 (0.51)	6.80 (0.41)	0.541	0.547†	0.524†
	P3b-L	Latency (ms)	317 (3.3)	317 (4.6)	318 (4.8)	321 (3.8)	324 (6.0)	319 (5.0)	0.165	0.354	0.320
	BPA	Accuracy (%)	96.5 (0.7)	96.8 (1.0)	96.1 (1.0)	89.8 (1.2)**	90.2 (1.6)	91.00 (1.6)	0.585	0.739††	0.694††
	MRT	Time (ms)	366 (8)	367 (12)	365 (12)	456 (11)**	456 (15)	443 (16)	0.767	0.755†††	0.778†††
Resting EEG	Delta-Abs	Power (μV^2/Hz)	109 (3.9)	105 (5.5)	112 (5.6)	138 (6.5)**	144 (9.6)	131 (8.6)	0.435	0.733††	0.715††
	Delta-Rel	n/a	0.329 (0.010)	0.327 (0.015)	0.332 (0.015)	0.341 (0.010)	0.349 (0.015)	0.333 (0.015)	0.085	0.802†††	0.806†††
	Theta-Abs	Power (μV^2/Hz)	58 (4.2)	59 (6.4)	58 (5.6)	74 (5.4)	73 (7.6)	76 (7.6)	0.255	0.781†††	0.892†††
	Theta-Rel	n/a	0.138 (0.005)	0.141 (0.008)	0.136 (0.007)	0.157 (0.007)	0.152 (0.008)	0.162 (0.010)	0.251	0.819†††	0.828†††
	Alpha1-Abs	Power (μV^2/Hz)	112 (11.0)	106 (14.6)	117 (16.4)	114 (10.0)	115 (15.3)	115 (13.1)	0.022	0.893†††	0.911†††
	Alpha1-Rel	n/a	0.221 (0.011)	0.217 (0.015)	0.225 (0.017)	0.221 (0.010)	0.218 (0.013)	0.225 (0.014)	0.022	0.883†††	0.878†††
	Alpha2-Abs	Power (μV^2/Hz)	44 (3.4)	45 (5.2)	42 (4.4)	47 (4.3)	48 (5.8)	46 (6.3)	0.076	0.769†††	0.882†††
	Alpha2-Rel	n/a	0.107 (0.006)	0.109 (0.008)	0.105 (0.008)	0.102 (0.006)	0.106 (0.009)	0.100 (0.008)	0.076	0.783†††	0.830†††
	Beta1-Abs	Power (μV^2/Hz)	37 (1.7)	36 (2.5)	37 (2.1)	38 (2.2)	38 (2.9)	38 (3.3)	0.266	0.796†††	0.692††
	Beta1-Rel	n/a	0.104 (0.003)	0.103 (0.004)	0.105 (0.005)	0.090 (0.003)*	0.089 (0.004)	0.090 (0.004)	0.398	0.896†††	0.699††
	Beta2-Abs	Power (μV^2/Hz)	12 (0.8)	12 (1.4)	11 (0.7)	11 (0.9)	11 (1.1)	12 (1.4)	0.005	0.325	0.781†††
	Beta2-Rel	n/a	0.033 (0.001)	0.034 (0.002)	0.032 (0.002)	0.028 (0.001)*	0.027 (0.001)	0.028 (0.002)	0.349	0.584†	0.831†††
	Beta3-Abs	Power (μV^2/Hz)	22 (1.3)	23 (1.4)	22 (1.6)	25 (2.2)	23 (2.0)	26 (4.0)	0.094	0.566†	0.397
	Beta3-Rel	n/a	0.067 (0.003)	0.069 (0.005)	0.065 (0.005)	0.060 (0.003)	0.069 (0.005)	0.065 (0.005)	0.348	0.735†††	0.638††
	TotalPow	Power (μV^2/Hz)	394 (19.5)	387 (28.6)	401 (26.6	449 (21)	453 (30.8)	445 (28.8)	0.214	0.864†††	0.885†††
	Gamma-Abs	Power (μV^2/Hz)	12 (1.0)	13 (1.6)	12 (1.3)	17 (2.6)	15 (1.8)	19 (4.8)	0.185	0.449†	0.278
	AlphaPeak	Frequency (Hz)	9.71 (0.09)	9.74 (0.09)	9.68 (0.13)	9.44 (0.10)	9.44 (0.14)	9.43 (0.14)	0.250	0.824†††	0.775†††
	SlowWave	n/a	0.98 (0.07)	0.97 (0.09)	0.99 (0.1)	0.91 (0.06)	0.88 (0.07)	0.94 (0.09)	0.081	0.846†††	0.797†††
	Theta/Beta	Ratio	1.23 (0.07)	1.24 (0.10)	1.22 (0.09)	1.71 (0.18)**	1.71 (0.17)	1.76 (0.17)	0.405	0.760†††	0.831†††

Abbreviations: FD-MMN = Frequency Deviant MMN; DD-MMN = Duration Deviant MMN; ASSR = Auditory Steady-State Response; HV_{AVG} = average of HV_{BL} and HV_{RT}; HV_{BL} = mean (±SEM) of all HV baseline tests; HV_{RT} = mean (±SEM) of all HV retests; SZ_{AVG} = average of SZ_{BL} and SZ_{RT}; SZ_{BL} = mean (±SEM) of all SZ baseline tests; SZ_{RT} = mean (±SEM) of all SZ retests; ICC = Interclass Correlation Coefficient.

[1] $^*p < 0.05$ and $^{**}p < 0.01$ compared to HV_{AVG}.

[2] † = Fair; †† = Good; ††† = Excellent (Cicchetti and Sparrow, 1981)

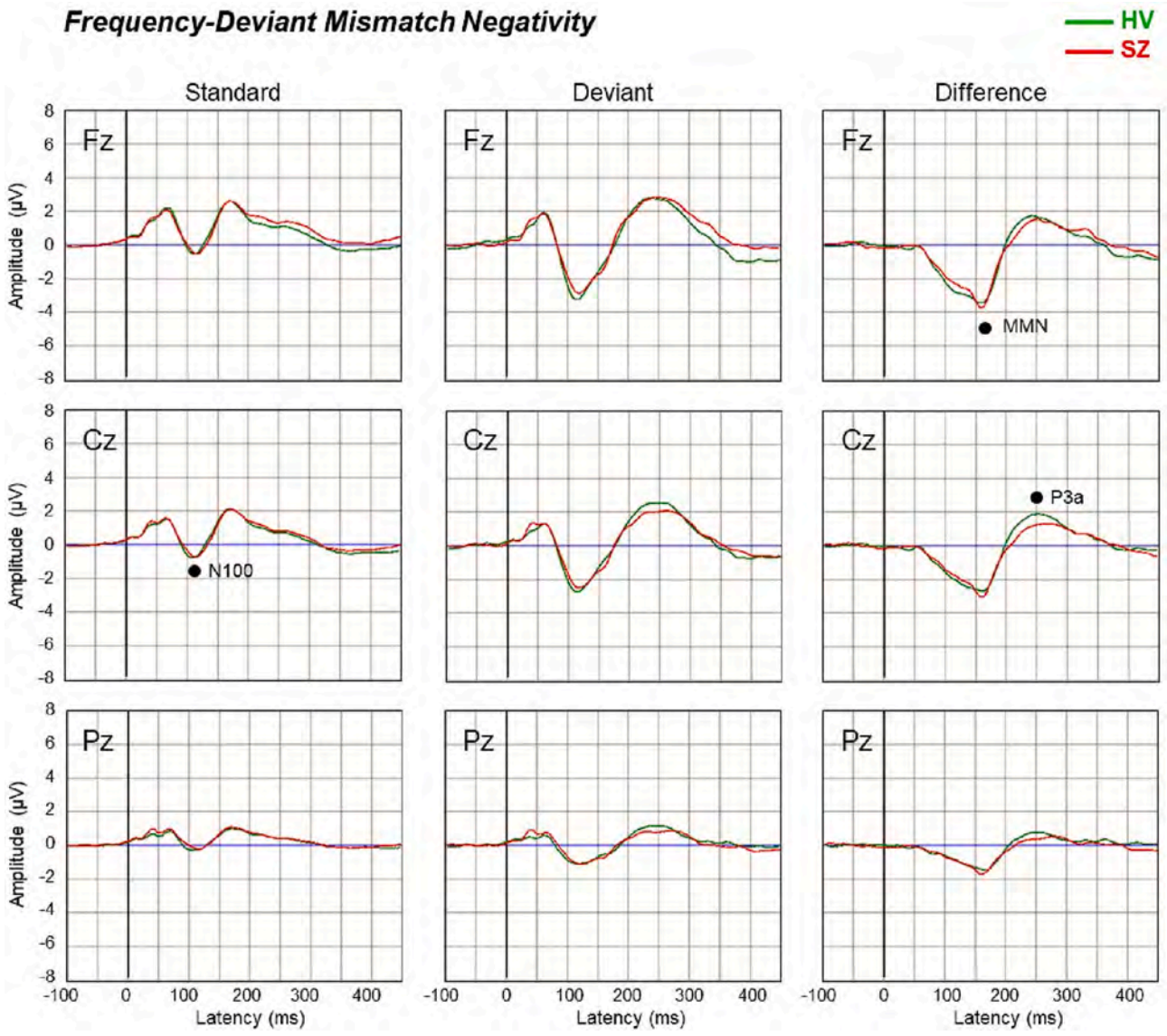

Fig. 2. Grand average and grand difference waves from midline electrodes for the frequency-deviant mismatch negativity paradigm.

analysis methods do not seem to affect findings of ASSR impairments in SZ (Thuné et al., 2016). Using a well-established preclinical model of acute NMDA hypofunction, Sivarao and colleagues have shown that EP is highly-sensitive to the level of NMDA receptor hypofunction (Sivarao et al., 2016). Thus, differences in EP deficits could at least in part reflect differences in the level of NMDAR receptor hypofunction across studied populations.

In the active oddball paradigm P3b amplitude was significantly decreased in SZ subjects. P3b deficits in SZ have been widely reported throughout the course of illness (see Onitsuka et al., 2013 for a review), and manifest since the early stages of the disease (Hamilton et al., 2019). In our study, P3b amplitude was also decreased at Retest when compared with Baseline regardless of subject group, suggesting possible habituation of the brain response to target deviants across repeated testing. Subjects with SZ showed longer reaction time and lower button press accuracy in the behavioral task associated with the active oddball. A delay in reaction time is already present early in SZ (Hamilton et al., 2019). Luck and colleagues have proposed that such deficit is the consequence of impairments in the response selection that lies between stimulus evaluation and response initiation (Luck et al., 2009).

Finally, for the eyes-closed resting EEG, a large published literature reports power increases across lower frequencies and decreases across higher frequencies in patients with SZ (for review, see Newson and Thiagarajan, 2019). Consistent with those findings, our results show an increase in Delta absolute power, a decrease in Beta1 and Beta2 relative power, and a significantly higher Theta/Beta ratio in the SZ group.

4.3. Correlations with functional assessments

A subset of ERP/QEEG endpoints that showed a deficit in SZ patients also correlated with functional assessments. Significant correlations were found with cognitive domains including attention and speed of information processing, working memory, verbal fluency, and functional capacity (Keefe et al., 2004, 2016), as well as the PANSS Positive and Negative Symptoms subscales. Correlations with functional assessments further underscore the utility of ERP/QEEG measures, suggesting that they might provide insight on the severity of the impairment in SZ at the subject level while tracking responses to treatments that improve function in this population.

Correlations for P3a amplitude from the frequency-deviant and the duration-deviant MMN paradigms did not overlap, suggesting again that these deviant types might engage distinct brain circuits (Curtis et al., 2021; Lee et al., 2017).

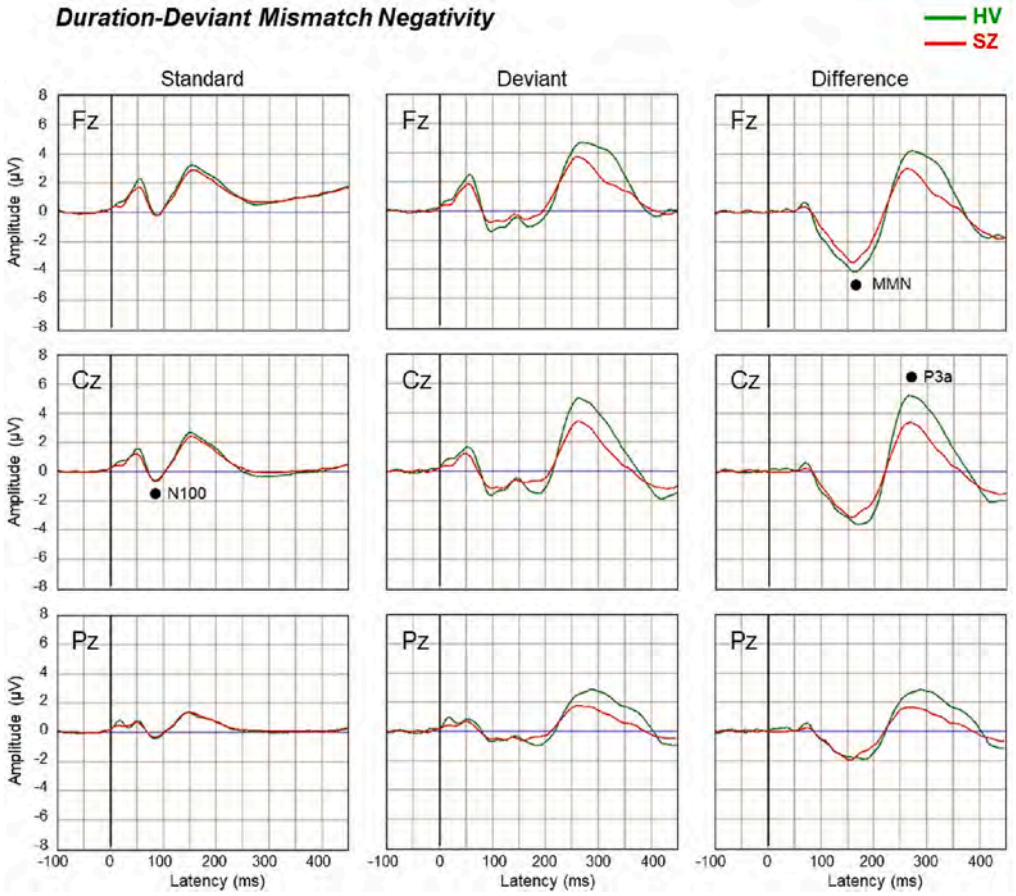

Fig. 3. Grand average and grand difference waves from midline electrodes for the duration-deviant mismatch negativity paradigm.

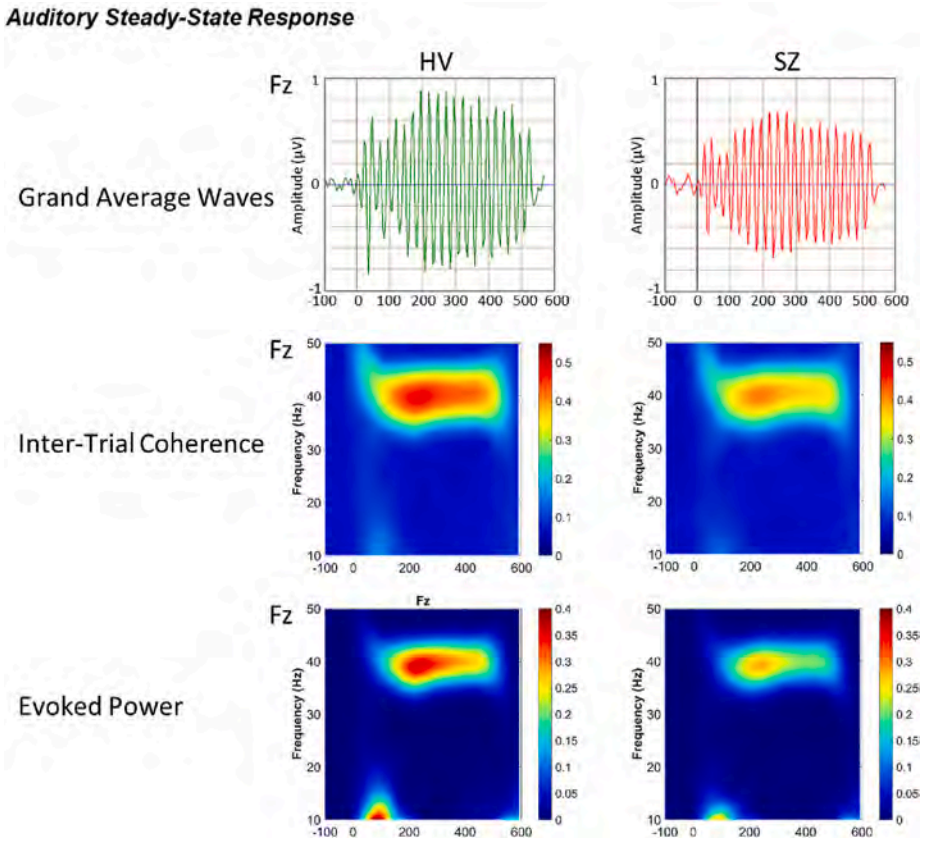

Fig. 4. Grand average waves (Top), and frequency-vs-time plots for ITC (Middle) and EP (Bottom) for the 40 Hz ASSR. Data is shown at the FZ electrode.

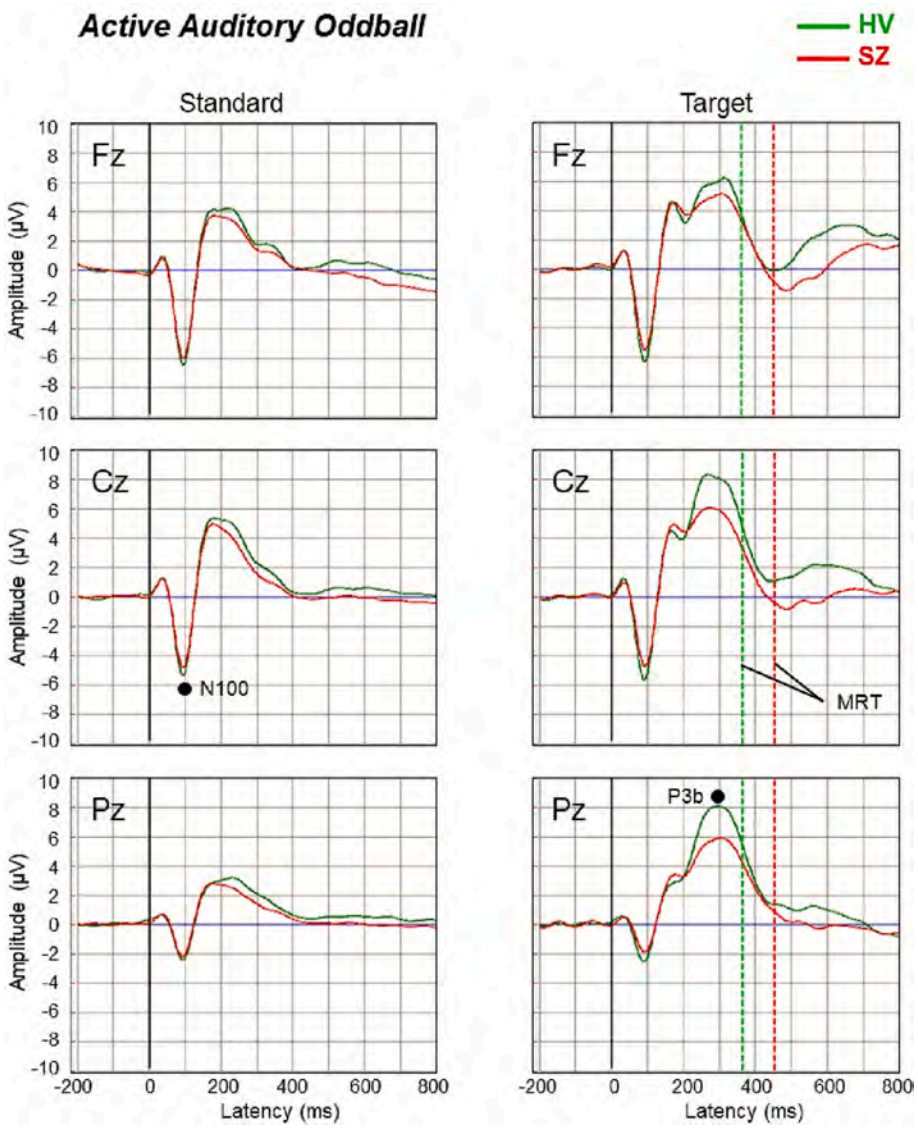

Fig. 5. Grand average waves from midline electrodes for the active auditory oddball. Dotted lines show median reaction time (MRT).

4.4. Study limitations

All patients were medicated. Thus, the study cannot distinguish effects of medication from those of the illness. Nevertheless, the critical issues were the effect-size and test-retest reliability of our measures in a subject sample that is likely representative of subjects who would participate in clinical trials of new cognition-targeted therapies in SZ.

4.5. Conclusions

The current study reports findings from a precompetitive, industry-led, collaborative research program. Our findings match published results from top academic labs, and show that complex ERP/EEG testing sessions can reliably be performed across clinical and commercial trial sites. The metrics reported on test-retest reliability can be leveraged for accurate power analyses in future interventional trials. The use of standardized equipment and protocols will allow scalability and ensure high data reproducibility across studies. The implementation of a fully automated ERP/EEG data analysis pipeline will facilitate ongoing study monitoring and adaptive study designs that can improve the likelihood of success and reduce costs.

Funding source

This research was solely funded by the ERP Biomarker Qualification Consortium. List of the Consortium members is available at https://erpbiomarkers.org/.

CRediT authorship contribution statement

All authors contributed to study conception and design, analysis and interpretation of results, and draft manuscript preparation. All authors approved the final version of the manuscript.

Declaration of competing interest

M. Cecchi is an employee of Cognision. M. Adachi and S. Honda are employees of Astellas Pharma. A. Basile. M.J. Marino, and J.M. Uslaner are employees of Merck & Co. D.L. Buhl is an employee and shareholder of Takeda Pharmaceuticals. H. Chadchankar and M. Rotte are employees of Novartis. S. Christensen is an employee of Lundbeck. E. Christian is an employee of Gilgamesh Pharmaceuticals. J. Doherty and B. Farley are employees and shareholders of Sage Therapeutics. K.C. Fadem is an employee and shareholder of Cognision. M.S. Forman is an employee of Passage Bio. J. Johannesen, P. O'Donnell, and M.C. Quirk are employees

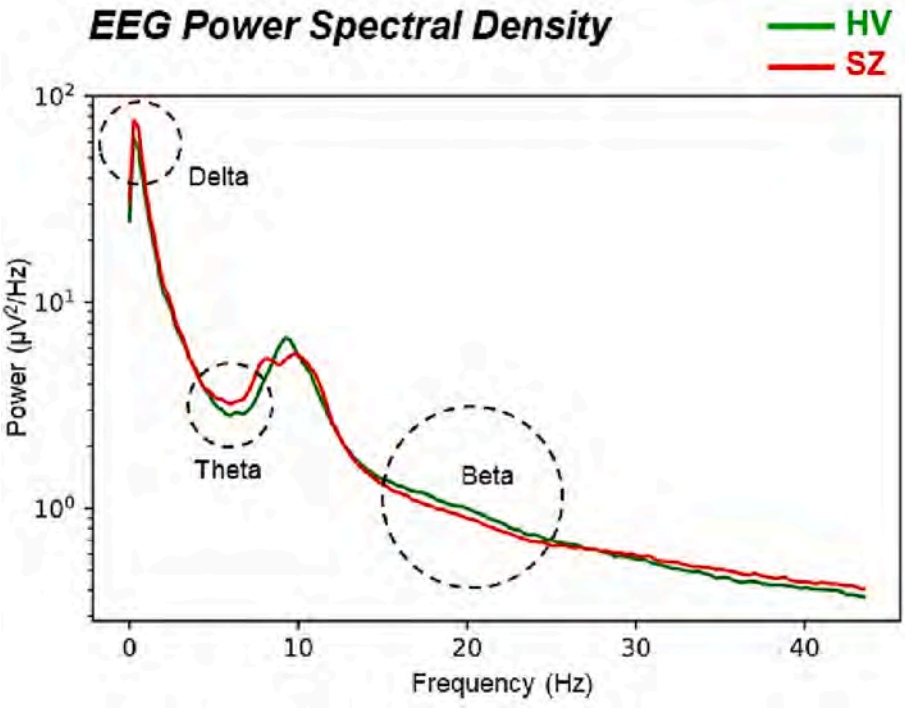

Fig. 6. Power spectral density for HV and SZ subjects from the eyes-closed resting state EEG test. Data is average of all electrodes.

Table 6
EEG/ERP parameter correlations with standard psychometric measures.

EEG/ERP Test	EEG/ERP Feature	Psychometric Measure	Correlation Coefficient
Frequency-Deviant MMN	P3a, Amplitude, Cz[1]	PANSS_PS	−0.320*
	P3a, Latency, Cz[1]	BAC_SC	−0.349*
	P3a, Latency, Cz[1]	VRFCAT_FP	0.304*
Duration-Deviant MMN	MMN, Amplitude, Fz[2]	BAC_SC	−0.386*
	P3a, Amplitude, Cz[2]	BAC_SC	0.340*
Auditory Steady-State Response	ITC400, Fz	BAC_DS	0.242*
	ITC400, Fz	BAC_VF	0.240*
Active Auditory Oddball	BPA	VRFCAT_AT	−0.277*
	MRT	PANSS_NS	0.254*
Resting EEG	Delta-Abs, Avg	PANSS_PS	0.249*
	Theta/Beta	BAC_VF	−0.221*
	Theta/Beta	VRFCAT_FP	0.305**

Abbreviations: PANSS_PS = Positive Symptoms Subscale; BAC_SC = Symbol Coding; VRFCAT_FP = Total Forced Progressions; BAC_DS = Digit Sequencing; BAC_VF = Verbal Fluency; VRFCAT_AT = Adjusted Total Time; PANSS_NS = Negative Symptoms Subscale.

[1] Frequency-deviant MMN paradigm.

[2] Duration-deviant MMN paradigm.

* $p < 0.05$.

** $p < 0.01$

of Sage Therapeutics. B.J. Kinon is an employee of Cyclerion Therapeutics. D. Klamer and C. Missling are employees and shareholders of Anavex Life Sciences Corp. T. Piser is an employee of Onsero Therapeutics; T. Piser has financial interest in the development of NMDA receptor modulators owned by Novartis Pharmaceuticals, and is an inventor on issued patents on some of these compounds. C.B. Puryear is an employee of Praxis Precision Medicines. C. Sanchez and D.G. Smith are employees and shareholders of Alkermes. D.C. Javitt has received consulting payments within the last 2 years from Autifony, Biogen, SK Life Sciences, Boehringer Ingelheim, and Biogen; he holds intellectual property rights for use of NMDA modulators in treatment of neuropsychiatric disorders, for parcel-guided TMS treatment of depression, and for EEG-based diagnosis of neuropsychiatric disorders; he also holds equity in Glytech, AASI, and NeuroRx. R.S.E. Keefe is the owner of VeraSci, a for-profit company that supports clinical trials for over 100 business entities, mostly pharmaceutical companies; VeraSci provided paid support for this study. D.P. Walling has received grant funding from Novartis, Janssen, Intracellular, Lyndra, Rovi, Otsuka, Alkermes, Cerevel, Abbvie, Allergan, Noven, Takeda, Indivior, Avanir, Lundbeck, Roche, Boehringer Ingelheim, Biogen, Sunovion and Acadia; he has served on Advisory Boards for Otsuka, Janssen, Biogen, Boehringer Ingelheim and Lyndra. L. Ereshefsky receives support from CenExel, which conducts research for most pharma, and is a performing site for NIH and Alzheimer's Foundation. Moreover, Follow the Molecule LLC receives consulting compensation from Bioxcel, Blackthorn, Neurocrine, Athira, Ceravance, Digestome, Immune Brain Check, Gilgamesh, and Karuna Therapeutics.

Acknowledgements

This research was funded by the ERP Biomarker Qualification Consortium. The Consortium members would like to thank all study participants for their contribution to the study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.schres.2023.02.018.

References

Addington, D., Addington, J., Schissel, B., 1990. A depression rating scale for schizophrenics. Schizophr. Res. 3, 247–251. https://doi.org/10.1016/0920-9964(90)90005-R.

Atkins, A.S., Tseng, T., Vaughan, A., Twamley, E.W., Harvey, P., Patterson, T., Narasimhan, M., Keefe, R.S.E., 2017. Validation of the tablet-administered Brief Assessment of Cognition (BAC App). Schizophr. Res. 181, 100–106. https://doi.org/10.1016/j.schres.2016.10.010.

Barbé, K., Pintelon, R., Schoukens, J., 2010. Welch method revisited: nonparametric power spectrum estimation via circular overlap. IEEE Trans. Signal Process. 58, 553–565. https://doi.org/10.1109/TSP.2009.2031724.

Cecchi, M., Moore, D.K., Sadowsky, C.H., Solomon, P.R., Doraiswamy, P.M., Smith, C.D., Jicha, G., Budson, A.E., Arnold, S.E., Fadem, K.C., 2015. A clinical trial to validate event-related potential markers of Alzheimer's disease in outpatient settings. Alzheimer's Dement. Diagn. Assess. Dis. Monit. 1, 387–394. https://doi.org/10.1016/j.dadm.2015.08.004.

Cicchetti, D.V., Sparrow, S.A., 1981. Developing criteria for establishing interrater reliability of specific items: applications to assessment of adaptive behavior. Am. J. Ment. Defic. 86, 127–137.

Clementz, B., Sweeney, J.A., Hamm, J.P., Ivleva, E.I., Ethridge, L.E., Pearlson, G.D., Keshavan, M.S., Tamminga, C.A., Hospital, H., 2017. Identification of distinct psychosis biotypes using brain-based. Biomarkers 173, 373–384. https://doi.org/10.1176/appi.ajp.2015.14091200.

Curtis, M.T., Coffman, B.A., Salisbury, D.F., 2021. Pitch and duration mismatch negativity are associated with distinct auditory cortex and inferior frontal cortex volumes in the first-episode schizophrenia spectrum. Schizophr. Bull. Open 2. https://doi.org/10.1093/SCHIZBULLOPEN/SGAB005.

Dondé, C., Martínez, A., Kantrowitz, J.T., Silipo, G., Dias, E.C., Patel, G.H., Sanchez-Peña, J., Corcoran, C.M., Medalia, A., Saperstein, A., Vail, B., Javitt, D., 2019. Bimodal distribution of tone-matching deficits indicates discrete pathophysiological entities within the syndrome of schizophrenia. Transl. Psychiatry 9. https://doi.org/10.1038/S41398-019-0557-8.

FDA, 2016. Collection of Race and Ethnicity Data in Clinical Trials: Guidance for Industry and Food and Drug Administration Staff. October 26, pp. 1–20.

Ford, J.M., Roach, B., Miller, R.M., Duncan, C.C., Hoffman, R.E., Mathalon, D.H., 2010. When it's time for a change: failures to track context in schizophrenia. Int. J. Psychophysiol. 78, 3–13. https://doi.org/10.1016/j.ijpsycho.2010.05.005.

Frodl, T., Meisenzahl, E.M., Müller, D., Greiner, J., Juckel, G., Leinsinger, G., Hahn, H., Möller, H.J., Hegerl, U., 2001. Corpus callosum and P300 in schizophrenia. Schizophr. Res. 49, 107–119. https://doi.org/10.1016/S0920-9964(00)00123-7.

Hamilton, H.K., Roach, B., Bachman, P.M., Belger, A., Carrion, R.E., Duncan, E., Johannesen, J.K., Light, G., Niznikiewicz, M.A., Addington, J., Bearden, C.E., Cadenhead, K.S., Cornblatt, B.A., McGlashan, T.H., Perkins, D.O., Seidman, L.J., Tsuang, M.T., Walker, E.F., Woods, S.W., Cannon, T.D., Mathalon, D.H., 2019. Association between P300 responses to auditory oddball stimuli and clinical outcomes in the psychosis risk syndrome. JAMA Psychiatry 76, 1187–1197. https://doi.org/10.1001/jamapsychiatry.2019.2135.

Hirano, Y., Oribe, N., Kanba, S., Onitsuka, T., Nestor, P.G., Spencer, K.M., 2015. Spontaneous gamma activity in schizophrenia. JAMA Psychiatry 72, 813–821. https://doi.org/10.1001/jamapsychiatry.2014.2642.

ICH Harmonised Tripartite Guideline, 2016. Guideline for good clinical practice E6(R2). ICH Harmon. Tripart. Guidel. 2016, 1–66.

Jasper, H., 1958. The ten twenty electrode system of the international federation. Electroencephalogr. Clin. Neurophysiol. 10, 371–375.

Javitt, D., Siegel, S.J., Spencer, K.M., Mathalon, D., Hong, L.E., Martinez, A., Ehlers, C.L., Abbas, A.I., Teichert, T., Lakatos, P., Womelsdorf, T., 2020. A roadmap for development of neuro-oscillations as translational biomarkers for treatment development in neuropsychopharmacology. Neuropsychopharmacology 45, 1411–1422. https://doi.org/10.1038/s41386-020-0697-9.

Jobert, M., Wilson, F.J., Ruigt, G.S.F., Brunovsky, M., Prichep, L.S., Drinkenburg, W., 2012. Guidelines for the recording and evaluation of pharmaco-EEG data in man: the International Pharmaco-EEG Society (IPEG). Neuropsychobiology 66, 201–220. https://doi.org/10.1159/000343478.

Kantrowitz, J.T., Epstein, M.L., Lee, M., Lehrfeld, N., Nolan, K.A., Shope, C., Petkova, E., Silipo, G., Javitt, D., 2017. Improvement in mismatch negativity generation during d-serine treatment in schizophrenia: correlation with symptoms. Schizophr. Res. 191, 70–79. https://doi.org/10.1016/j.schres.2017.02.027.

Kay, S.R., Fiszbein, A., Opler, L.A., 1987. The positive and negative syndrome scale (PANSS) for schizophrenia. Schizophr. Bull. 13, 261–276. https://doi.org/10.1093/schbul/13.2.261.

Keefe, R.S.E., Goldberg, T.E., Harvey, P.D., Gold, J.M., Poe, M.P., Coughenour, L., 2004. The brief assessment of cognition in schizophrenia: reliability, sensitivity, and comparison with a standard neurocognitive battery. Schizophr. Res. 68, 283–297. https://doi.org/10.1016/j.schres.2003.09.011.

Keefe, R.S.E., Davis, V.G., Atkins, A.S., Vaughan, A., Patterson, T., Narasimhan, M., Harvey, P.D., 2016. Validation of a computerized test of functional capacity. Schizophr. Res. 175, 90–96. https://doi.org/10.1016/j.schres.2016.03.038.

Kirihara, K., Rissling, A.J., Swerdlow, N.R., Braff, D.L., Light, G., 2012. Hierarchical organization of gamma and theta oscillatory dynamics in schizophrenia. Biol. Psychiatry 71, 873–880.

Lee, M., Sehatpour, P., Dias, E.C., Silipo, G.S., Kantrowitz, J.T., Martinez, A.M., Javitt, D., 2017. A tale of two sites: differential impairment of frequency and duration mismatch negativity across a primarily inpatient versus a primarily outpatient site in schizophrenia. Schizophr. Res. 191, 10–17. https://doi.org/10.1016/j.schres.2017.07.032.

Li, Z., Zheng, B., Deng, W., Liu, X., Zheng, Z., Li, T., 2013. Multi-components of evoked-brain potentials in deficit and nondeficit schizophrenia. Asia Pac. Psychiatry 5, 69–79. https://doi.org/10.1111/appy.12030.

Light, G., Hsu, J.L., Hsieh, M.H., Meyer-Gomes, K., Sprock, J., Swerdlow, N.R., Braff, D.L., 2006. Gamma band oscillations reveal neural network cortical coherence dysfunction in schizophrenia patients. Biol. Psychiatry 60, 1231–1240. https://doi.org/10.1016/j.biopsych.2006.03.055.

Light, G., Swerdlow, N.R., Thomas, M.L., Calkins, M.E., Green, M.F., Greenwood, T.A., Gur, R.E., Gur, R.C., Lazzeroni, L.C., Nuechterlein, K.H., Pela, M., Radant, A.D., Seidman, L.J., Sharp, R.F., Siever, L.J., Silverman, J.M., Sprock, J., Stone, W.S., Sugar, C.A., Tsuang, D.W., Tsuang, M.T., Braff, D.L., Turetsky, B.I., 2014. Validation of mismatch negativity and P3a for use in multi-site studies of schizophrenia: characterization of demographic, clinical, cognitive, and functional correlates in COGS-2. Schizophr. Res. 163, 63–72. https://doi.org/10.1016/j.schres.2014.09.042.

Light, G.A., Braff, D.L., 2005. Stability of mismatch negativity deficits and their relationship to functional impairments in chronic schizophrenia. Am. J. Psychiatry 162, 1741–1743. https://doi.org/10.1176/APPI.AJP.162.9.1741.

Luck, S.J., Kappenman, E.S., Fuller, R.L., Robinson, B., Summerfelt, A., Gold, J.M., 2009. Impaired response selection in schizophrenia: evidence from the P3 wave and the lateralized readiness potential. Psychophysiology 46, 776–786. https://doi.org/10.1111/j.1469-8986.2009.00817.x.

Luck, S.J., Mathalon, D., O'Donnell, B.F., Hämäläinen, M.S., Spencer, K.M., Javitt, D., Uhlhaas, P.J., 2011. A roadmap for the development and validation of event-related potential biomarkers in schizophrenia research. Biol. Psychiatry 70, 28–34. https://doi.org/10.1016/j.biopsych.2010.09.021.

Mathalon, D.H., Ford, J.M., Pfefferbaum, A., 2000. Trait and state aspects of p300 amplitude reduction in schizophrenia: a retrospective longitudinal study. Biol. Psychiatry 47, 434–449. https://doi.org/10.1016/S0006-3223(99)00277-2.

Nagai, T., Tada, M., Kirihara, K., Yahata, N., Hashimoto, R., Araki, T., Kasai, K., 2013. Auditory mismatch negativity and P3a in response to duration and frequency changes in the early stages of psychosis. Schizophr. Res. 150, 547–554. https://doi.org/10.1016/j.schres.2013.08.005.

Newson, J.J., Thiagarajan, T.C., 2019. EEG frequency bands in psychiatric disorders: a review of resting state studies. Front. Hum. Neurosci. 12, 1–24. https://doi.org/10.3389/fnhum.2018.00521.

O'Donnell, B.F., Vohs, J.L., Krishnan, G.P., Rass, O., Hetrick, W.P., Morzorati, S.L., 2013. The auditory steady-state response (ASSR): a translational biomarker for schizophrenia. Suppl. Clin. Neurophysiol. 62, 101–112. https://doi.org/10.1016/B978-0-7020-5307-8.00006-5.

O'Donnell, P., Rosen, L., Alexander, R., Murthy, V., Davies, C.H., Ratti, E., 2019. Strategies to address challenges in neuroscience drug discovery and development. Int. J. Neuropsychopharmacol. 22, 445–448. https://doi.org/10.1093/IJNP/PYZ027.

Onitsuka, T., Oribe, N., Nakamura, I., Kanba, S., 2013. Review of neurophysiological findings in patients with schizophrenia. Psychiatry Clin. Neurosci. 1–10 https://doi.org/10.1111/pcn.12090.

Roach, B., Mathalon, D., 2008. Event-related EEG time-frequency analysis: an overview of measures and an analysis of early gamma band phase locking in schizophrenia. Schizophr. Bull. 34, 907–926. https://doi.org/10.1093/schbul/sbn093.

Roach, B., Hamilton, H.K., Bachman, P., Belger, A., Carrión, R.E., Duncan, E., Johannesen, J.K., Kenney, J.G., Light, G., Niznikiewicz, M., Addington, J., Bearden, C.E., Owens, E.M., Cadenhead, K.S., Cannon, T.D., Cornblatt, B.A., McGlashan, T.H., Perkins, D.O., Seidman, L., Tsuang, M., Walker, E.F., Woods, S.W., Mathalon, D., 2020. Stability of mismatch negativity event-related potentials in a multisite study. Int. J. Methods Psychiatr. Res. 1–15 https://doi.org/10.1002/mpr.1819.

Salisbury, D.F., Kuroki, N., Kasai, K., Shenton, M.E., McCarley, R.W., 2007. Progressive and interrelated functional and structural evidence of post-onset brain reduction in schizophrenia. Arch. Gen. Psychiatry 64, 521–529. https://doi.org/10.1001/ARCHPSYC.64.5.521.

Sheehan, D.V., Lecrubier, Y., Sheehan, K.H., Amorim, P., Janavs, J., Weiller, E., Hergueta, T., Baker, R., Dunbar, G.C., 1998. The Mini-international neuropsychiatric interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J. Clin. Psychiatry 59 (Suppl. 2), 22–33 quiz 34-57.

Simpson, G.M., B, M., B, G., Angus, J.W.S., Pear, F.R.C., Mat, D.P., 1970. A rating scale for extrapyramidal side effects. Acta Psychiatr. Scand. Suppl. 212, 11–19. https://doi.org/10.1111/J.1600-0447.1970.TB02066.X.

Sivarao, D.V., Chen, P., Senapati, A., Yang, Y., Fernandes, A., Benitex, Y., Whiterock, V., Li, Y.-W., Ahlijanian, M.K., 2016. 40 Hz auditory steady-state response is a pharmacodynamic biomarker for cortical NMDA receptors. Neuropsychopharmacology 41, 2232–2240.

Thomas, M.L., Green, M.F., Hellemann, G., Sugar, C.A., Tarasenko, M., Calkins, M.E., Greenwood, T.A., Gur, R.E., Gur, R.C., Lazzeroni, L.C., Nuechterlein, K.H., Radant, A.D., Seidman, L.J., Shiluk, A.L., Siever, L.J., Silverman, J.M., Sprock, J., Stone, W.S., Swerdlow, N.R., Tsuang, D.W., Tsuang, M.T., Turetsky, B.I., Braff, D.L., Light, G., 2017. Modeling deficits from early auditory information processing to psychosocial functioning in schizophrenia supplemental content. JAMA Psychiatry 74, 37. https://doi.org/10.1001/jamapsychiatry.2016.2980.

Thuné, H., Recasens, M., Uhlhaas, P.J., 2016. The 40-Hz auditory steady-state response in patients with schizophrenia: a meta-analysis. JAMA Psychiatry 73, 1145–1153.

Turetsky, B.I., Calkins, M.E., Light, G.A., Olincy, A., Radant, A.D., Swerdlow, N.R., 2007. Neurophysiological endophenotypes of schizophrenia: the viability of selected candidate measures. Schizophr. Bull. 33, 69–94. https://doi.org/10.1093/schbul/sbl060.