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

https://escholarship.org/uc/item/8hz3k553

Schizophrenia Research, 163(1-3)

0920-9964

Light, GA
Swerdlow, NR
Thomas, ML
et al.

2015

10.1016/j.schres.2014.09.042

Peer reviewed
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

Gregory A. Light a,b,⁎, Neal R. Swerdlow a, Michael L. Thomas a, Monica E. Calkins c, Michael F. Green d,e, Tiffany A. Greenwood a, Raquel E. Gur c, Ruben C. Gur c, Laura C. Lazzeroni f, Keith H. Nuechterlein d, Marlena Pela a, Allen D. Radant g,h, Larry J. Seidman i,j, Richard F. Sharp a, Larry J. Siever k,l, Jeremy M. Silverman k,l, Joyce Sprock a, William S. Stone m, Catherine A. Sugar m, Debby W. Tsuang g,h, Ming T. Tsuang a,n,o, David L. Braff a, Bruce I. Turetsky c

a Department of Psychiatry, University of California San Diego, La Jolla, CA, United States
b VISN-22 Mental Illness, Research, Education and Clinical Center (MIRECC), VA San Diego Healthcare System, United States
c Department of Psychiatry, University of Pennsylvania, Philadelphia, PA, United States
d Department of Psychiatry and Biobehavioral Sciences, University of California Los Angeles, Los Angeles, CA, United States
e VA Greater Los Angeles Healthcare System, Los Angeles, CA, United States
f Department of Psychiatry and Behavioral Sciences, Stanford University, Stanford, CA, United States
g Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, WA, United States
h VA Puget Sound Health Care System, Seattle, WA, United States
i Department of Psychiatry, Harvard Medical School, Boston, MA, United States
j Massachusetts Mental Health Center Public Psychiatry Division of the Beth Israel Deaconess Medical Center, Boston, MA, United States
k Department of Psychiatry, The Mount Sinai School of Medicine, New York, NY, United States
l James J. Peters VA Medical Center, New York, NY, United States
m Department of Biostatistics, University of California Los Angeles School of Public Health, Los Angeles, CA, United States
n Center for Behavioral Genomics, and Institute for Genomic Medicine, University of California San Diego, La Jolla, CA, United States
o Harvard Institute of Psychiatric Epidemiology and Genetics, Boston, MA, United States

Abstract

Mismatch negativity (MMN) and P3a are auditory event-related potential (ERP) components that show robust deficits in schizophrenia (SZ) patients and exhibit qualities of endophenotypes, including substantial heritability, test–retest reliability, and trait-like stability. These measures also fulfill criteria for use as cognition and function-linked biomarkers in outcome studies, but have not yet been validated for use in large-scale multi-site clinical studies. This study tested the feasibility of adding MMN and P3a to the ongoing Consortium on the Genetics of Schizophrenia (COGS) study. The extent to which demographic, clinical, cognitive, and functional characteristics contribute to variability in MMN and P3a amplitudes was also examined. Participants (HCS n = 824, SZ n = 966) underwent testing at 5 geographically distributed COGS laboratories. Valid ERP recordings were obtained from 91% of HCS and 91% of SZ patients. Highly significant MMN (d = 0.96) and P3a (d = 0.93) amplitude reductions were observed in SZ patients, comparable in magnitude to those observed in single-lab studies with no appreciable differences across laboratories. Demographic characteristics accounted for 26% and 18% of the variance in MMN and P3a amplitudes, respectively. Significant relationships were observed among demographically-adjusted MMN and P3a measures and medication status as well as several clinical, cognitive, and functional characteristics of the SZ patients. This study demonstrates that MMN and P3a ERP biomarkers can be feasibly used in multi-site clinical studies. As with many clinical tests of brain function, demographic factors contribute to MMN and P3a amplitudes and should be carefully considered in future biomarker-informed clinical studies.

1. Introduction

There is compelling evidence that sensory processing impairments contribute to the cognitive and psychosocial dysfunction affecting the majority of schizophrenia (SZ) patients (e.g., Braff and Light, 2004;...
MMN also meets the criteria for use as an endophenotype in geno-
phrenia Network), to our knowledge there are no published studies of
large cohort of participants tested across the COGS-2 laborato-
the characterization of demographic factors associated with MMN and P3a amplitudes including age, sex, race, medication type, and smoking status. Lastly, we tested for significant bi-
variate relationships among MMN and P3a with measures of clinical, cognitive, and psychosocial functioning after accounting for identified
demographic factors.

2. Methods

2.1. Participants

Participants included 1790 (HCS n = 824, SZ n = 966) subjects that were recruited and tested at the 5 COGS-2 test sites: University of California San Diego (UCSD), University of California Los Angeles (UCLA), University of Washington (UW), University of Pennsylvania (Penn), and Mount Sinai School of Medicine (MSSM). All participants were assessed on their capacity to provide informed consent. After subjects were given a detailed description of their participation in the study, written consent was obtained via methods approved by the local human research protection committees at each testing site prior to participation (UCSD Protocol #080425). Exclusionary factors included evidence of Axis I psychiatric and neurological disorders other than schizophrenia, head injury, stroke, substance abuse (except tobacco) or a history of psychotic disorders in first degree relatives of HCS as determined by the Family Interview for Genetic Studies (Maxwell, 1992).

MMN and P3a are informative probes of the neural substrates of sensory processing abnormalities. These measures are supported by a distributed network of frontaltemporal cortical sources underlying pas-
receptor functioning.

The temporal window of early information processing reflected by MMN and P3a appears to be a critical transitional zone from sensory-based processing to the engagement of higher attentional neural net-
works necessary for cognitive and psychosocial functioning (Rissling et al., 2013a). Indeed, previous studies have demonstrated that MMN and P3a are each significantly correlated with distinct domains of cogni-
tive (Baldeweg et al., 2004; Näätänen et al., 2011; Light et al., 2007; Kawakubo et al., 2006) and psychosocial functioning (Light and Braff 2005a, 2005b; Kawakubo et al., 2007; Wynn et al., 2010; Rasser et al., 2011). These components also exhibit utility as repeated measures over short and long (e.g., 12-month) retest intervals in both healthy subjects and SZ patients (Light et al., 2012; ICCs ≈ 0.90). Reliability coefficients and effect sizes of deficits in SZ for both MMN and P3a are comparable to or even exceed those obtained from standard neuropsy-
chological tests commonly used in SZ research (Light et al., 2012; Light and Braff, 2005a,b). This collection of attributes has contributed to the view of MMN as a “breakthrough biomarker” (Light and Näätänen, 2013) that is “translatable” (Nagai et al., 2013a, 2013b) and potentially very important (Belger et al., 2012) in neuropsychiatry.

MMN also meets the criteria for use as an endophenotype in geno-
mic studies as it is highly heritable (Hall et al., 2006, 2009; Price et al., 2006), independent of fluctuations of clinical state and symptoms (Shinozaki et al., 2002; Light et al., 2012), and present in individuals at genetic risk for developing schizophrenia (Schreiber et al., 1992; Jessen et al., 2001; Michie et al., 2002; Baker et al., 2005; Atkinson et al., 2012). The NIMH Consortium on the Genetics of Schizophrenia (COGS) has pursued multi-site genetic studies of schizophrenia and related endophenotypes, first in a family study (Calkins et al., 2007; Light et al., 2014) of SZ probands, unaffected family members and Healthy Comparison Subjects (HCS) and then more recently in a larger case–control study of SZ patients and unrelated HCS (COGS-2) described in this issue. Given the growing importance of MMN in schizophrenia research, this measure was added to COGS-2 in study years 2–4.

Sample size demands of genetic studies often require the use of mul-
tiple data collection sites. As we have previously noted (Swerdlov et al., 2014), this approach may present challenges for studies of complex phenotypes like MMN because of potential differences in laboratory conditions or sample characteristics across sites that introduce uncon-
trolled variance into experimental measures. On the other hand, by test-
ing a more heterogeneous sample, multi-site studies increase the likelihood that findings will be generalizable rather than site-specific. Although there are now other consortium studies using ERPs in special-
ized academic laboratories with expertise in multi-sensor recordings (e.g., North American Prodromal Longitudinal Study, Bipolar & Schizo-
phrenia Network), to our knowledge there are no published studies of the feasibility and fidelity of recording MMN and P3a using a simple, 2-channel system tested in laboratories that do not specialize in EEG ac-
quisition. In this study, we first examined the “yield” of usable data ob-
tained from the large cohort of participants tested across the COGS-2 laborato-
ries reproduce

2.2. Stimuli and procedures

A duration-deviant auditory oddball paradigm was employed follow-
ing our established procedures (Light and Braff, 2005a; Light et al., 2007, 2010, 2012; Kiang et al., 2009; Rissling et al., 2010, 2012, 2013a); see Fig. 1. Subjects were presented with binaural tones (1-kHz, 85-dB, with 1-ms rise/fall, stimulus onset-to-onset asynchrony 500 ms) via insert earphones (Aearo Company Auditory Systems, Indianapolis, IN; Model

The temporal window of early information processing reflected by MMN and P3a appears to be a critical transitional zone from sensory-based processing to the engagement of higher attentional neural net-
works necessary for cognitive and psychosocial functioning (Rissling et al., 2013a). Indeed, previous studies have demonstrated that MMN and P3a are each significantly correlated with distinct domains of cogni-
tive (Balde...
3A). Standard \((p = 0.90, 50\text{-ms duration})\) and Deviant \((p = 0.10, 100\text{-ms duration})\) tones were presented in pseudorandom order with a minimum of 6 Standard stimuli presented between each Deviant stimulus. During the approximately 20-min session, participants watched a silent cartoon video. Participants were instructed to attend to the video as they might be asked to answer questions about it at the end of the session.

2.3. Electroencephalographic (EEG) recording

A 2-channel EEG system with a pre-set MMN module was used for stimulus presentation and EEG recording (ERPlab, San Diego Instruments, San Diego, CA). One channel recorded EEG activity at the vertex (Cz) referenced to the left mastoid process (full scale setting 0.1, bandpass filter settings 0.5–100 kHz). A second channel recorded eye movement (EOG) activity from electrodes placed midsuperior and lateral to the right orbit (full scale setting 0.25, bandpass filter settings 0.5–100 kHz). A ground electrode was placed on the right mastoid. All electrode impedances were below 5 kΩ. Subjects were seated comfortably in front of a computer monitor and directed to fixate their gaze on the center of the screen. A hearing test was conducted to ensure <=40 dB hearing threshold bilaterally to 1000 kHz tones.

2.4. EEG data processing

Continuous EEG data from all 5 COGS-2 sites were uploaded to a centralized database. Quality assurance data review and analysis was then conducted by a single investigator (GL) who was blind to all demographic and diagnostic information. EEG data were processed using Brain Vision Analyzer 2.0 (Brain Products GmbH). Data were digitally filtered between 0.1 and 30 kHz (24 dB/oct) and eye movement artifact was removed using an automated correction algorithm. Intervals with additional EEG artifact (activity exceeding +/- 50 μV) at the Cz electrode were excluded from further analysis. The remaining trials were then sorted and combined to form separate average ERP waveforms for the deviant and standard tone conditions. MMN and P3a difference waveforms were generated by subtracting ERPs in response to standard tones from the ERPs generated in response to the Deviant stimuli. These difference waves were baseline corrected relative to the 100 ms pre-stimulus interval and visually inspected to determine the presence or absence of reliably identifiable ERP components. A highly conservative and stringent approach to data inclusion was employed. Data without an unambiguously response to the standard tone, or a reliably identifiable MMN or P3a response to the difference waves were excluded from further analyses. MMN and P3a amplitudes were respectively quantified as the mean amplitude from 135 to 205 ms and from 250 to 300 ms time windows.

2.5. Assessment of functional capacity

Patients’ functional capacity was assessed with the abbreviated version of the UCSD Performance Based Skills Assessment (UPSA-B; Patterson et al., 2001). The UPSA directly measures functional skills, using standardized tasks that are commonly encountered in everyday situations and considered necessary for independent community living including: financial management and communication skills.

2.6. Assessment of psychosocial functional status

The Scale of Functioning was used to assess psychosocial functional status in domains of independent living, social, and instrumental functioning (Rapaport et al., 1996). The Role Functioning Scale (RFS; McPheeters, 1984) assesses working productivity, independent living/self care, family network relationships, and immediate social network relationship (close friends, spouse).

2.7. Analyses

A series of multiple linear regression models was separately applied using MMN and P3a as dependent variables. A strategy of comparing incremental fit of variables or variable sets was employed (Cohen et al., 2003). Four levels were defined as follows: 1) demographic variables (age, sex, and race); 2) group (patient status); 3) test site; and 4) group-by-site interaction. Each level was tested for significant
improvement in model fit, in order, with an alpha set to .05 with the additional constraint that the step accounts for a minimum of additional 5% of variance in the dependent variable. Only variables in sets that led to significant and meaningful improvement in model fit were retained (although the interaction models tested always included all relevant main effects). Finally, we tested the significance of select regression parameters only for those variables in sets shown to significantly improve model fit. All analyses were conducted using R (R Development Core Team, 2011).

3. Results

The total number of included participants for whom ERP data were uploaded to the COGS-2 database was 1790 (HCS n = 824,
SZ n = 966); 91% of participants in each group (HCS n = 753, SZ n = 877) were deemed to have data of sufficient quality to permit further analysis. Consistent with previous studies conducted in single laboratories, significant large effect size deficits in MMN (d = 0.96) and P3a (d = 0.93) were observed in SZ patients. Individual data for each of the 1630 participants and group waveform averages are shown in Fig. 2. Table 1 reports the clinical and demographic characteristics of the current samples (for an expanded table of the complete sample broken down by site see Swerdlow et al. (2015)). Significant group differences were detected in age (patients older than controls), sex (more male patients), self-identified racial composition, and smoking (more patients smoked). As expected, SZ and HCS also differed on education and Mini-Mental State Exam scores.

For MMN, the demographic characteristics of age, sex, and race accounted for 26% of variance, $F(4,1625) = 140.10, p < 0.001$ (Figs. 3 and 4). The addition of group significantly improved model fit ($p < 0.05; \Delta R^2 = .07$). The addition of site did not significantly improve model fit ($p = .77$; Fig. 5). The group-by-site interaction was statistically significant; however, the effect size was trivially small, accounting for only an additional 1% of variance in MMN ($p < 0.001; \Delta R^2 = .01$). In the model predicting MMN from demographic variables and group [$F(5,1624) = 155.10, p < 0.001, R^2 = .32$], age had the strongest demographic effect ($b = 0.03, p < .001$, $R_{\text{partial}}^2 = .11$). The effect of group (SZ) remained highly significant ($b = 0.65; p < .001; R_{\text{partial}}^2 = .08$) even after covarying demographic variables.

For P3a, demographic characteristics of age, sex, and race accounted for 18% of variance $F(4,1625) = 91.11, p < 0.001$ (Figs. 3 and 4). The addition of group significantly improved model fit ($p < 0.001; \Delta R^2 = .08$). The addition of site was statistically significant given the large N; however, the effect was trivially small, accounting for only 1% of additional variance in P3a ($p = .001; \Delta R^2 = .01$; Fig. 5). The addition of the group-by-site interaction term did not significantly improve model fit ($p = .36$). In the model predicting P3a from demographic variables and group [$F(5,1624) = 114.30, p < 0.001, R^2 = .26$], age had the

Table 1
Clinical and demographic characteristics of controls and patients.

<table>
<thead>
<tr>
<th></th>
<th>Controls (HCS)</th>
<th>Patients (SZ)</th>
<th>p</th>
<th>E.S.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample size</td>
<td>753</td>
<td>877</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age M (SD)</td>
<td>38.63 (12.00)</td>
<td>46.25 (11.23)</td>
<td>&lt;.001</td>
<td>.30</td>
</tr>
<tr>
<td>Education M (SD)</td>
<td>14.99 (2.20)</td>
<td>12.63 (2.15)</td>
<td>&lt;.001</td>
<td>.48</td>
</tr>
<tr>
<td>Male</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian n (%)</td>
<td>159 (21%)</td>
<td>162 (18%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>African American n</td>
<td>638 (58%)</td>
<td>380 (43%)</td>
<td>&lt;.001</td>
<td>.19</td>
</tr>
<tr>
<td>Other n (%)</td>
<td>156 (21%)</td>
<td>335 (38%)</td>
<td>.016</td>
<td>.09</td>
</tr>
<tr>
<td>Smoking n (%)</td>
<td>84 (11%)</td>
<td>464 (53%)</td>
<td>&lt;.001</td>
<td>.44</td>
</tr>
<tr>
<td>Site</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCSD</td>
<td>179 (24%)</td>
<td>252 (29%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCLA</td>
<td>186 (25%)</td>
<td>200 (23%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UW</td>
<td>159 (21%)</td>
<td>138 (16%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PENN</td>
<td>153 (20%)</td>
<td>202 (23%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSSM</td>
<td>76 (10%)</td>
<td>85 (10%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MMSE M (SD)</td>
<td>33.61 (1.69)</td>
<td>31.12 (3.31)</td>
<td>&lt;.001</td>
<td>.42</td>
</tr>
<tr>
<td>SAPS M (SD)</td>
<td>6.88 (4.09)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SANS M (SD)</td>
<td>11.64 (5.36)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age of onset M (SD)</td>
<td>22.52 (7.23)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. UCSD = University of California, San Diego; UCLA = University of California, Los Angeles; MSSM = Mount Sinai School of Medicine; PENN = University of Pennsylvania; UW = University of Washington; MMSE = Mini-Mental State Exam; SAPS = Scale for the Assessment of Positive Symptoms; SANS = Scale for the Assessment of Negative Symptoms. Effect sizes (E.S.) are reported as point-biserial correlation coefficients for continuous (x) variables and Cramer’s V (or simply phi) for categorical (x) variables.
The strongest demographic effect \( (b = -0.03, p < .001, R^2_{\text{partial}} = .04) \). The effect of group (SZ) remained highly significant \( (b = -0.97; p < .001; R^2_{\text{partial}} = .091) \) even after covarying demographic variables.

For both MMN and P3a, older age, male sex, African American race, and SZ all predicted significantly smaller amplitudes (i.e., less negative MMN and less positive P3a). Notably, impedance values were also higher for African Americans \( (M = 4.12; SE = .10) \) compared to Caucasians \( (M = 3.21; SE = .07) \).

This large sample allowed for secondary analyses within the SZ group only, including comparisons of unmedicated patients with those treated with different classes of antipsychotics, with vs. without anticholinergic treatment, and smokers vs. non-smokers. For MMN, the addition of medication type significantly improved model fit over demographic variables (age, sex, and race) alone. Compared to no medication, 1st generation antipsychotics \( (p = .035) \) and the combination of both 1st and 2nd generation antipsychotics \( (p = .001) \) predicted smaller (i.e., less negative) MMN amplitudes. In contrast, unmedicated patients did not differ from patients receiving 2nd generation antipsychotics \( (p = .129) \). For P3a, 1st generation antipsychotics \( (p = .002) \), 2nd generation antipsychotics \( (p < .001) \), and the combination of both \( (p = .001) \) all predicted smaller (less positive) amplitudes. Fig. 6 plots the MMN and P3a in SZ patients by medication type. MMN and P3a amplitudes were also significantly smaller in patients who were smokers vs. non-smokers (Fig. 8). HCS are included in Figs. 6–8 for reference, but were not included in the regression analyses.

Finally, Table 2 reports the correlations between demographically corrected MMN and P3a values and several additional clinical variables of interest. Significant correlations were observed among MMN and age of illness onset, positive and negative symptoms, global cognitive function (MMSE), performance on tasks necessary for independent functioning (UPSA), psychosocial functional status (SOF), and Role Functioning (RFS). P3a was significantly associated with age of illness onset and psychosocial functional status.

### 4. Discussion

MMN and P3a have emerged as promising biomarkers for understanding and treating psychotic disorders (Braff and Light, 2004; Javitt, 2009; Belger et al., 2012; Light and Näätänen, 2013; Nagai et al., 2013a; Light and Swerdlow, 2014; Perez et al., 2014a). These measures

---

1 An individual was classified as a smoker if they reported smoking at least 1 cigarette per day. The effect of smoking on MMN and P3a was also examined in the initial regression models. For MMN neither smoker nor group by smoker significantly improved the model fit. When the smoker variable was added to the regression model predicting P3a from demographic variables and the group status variable, the improvement in fit was trivial, accounting for less than 1% of added variance in MMN \( (p < .01; \Delta R^2 = .003) \). Adding a smoker by group interaction term did not significantly improve model fit compared to the model with a main effect alone. A similar, if not weaker, pattern of results was obtained when number of cigarettes smoked per day was used.
have already undergone extensive psychometric validation (Kiang et al., 2009; Light et al., 2012; Rissling et al., 2012; Takahashi et al., 2012). MMN has demonstrated utility for forecasting the duration of time to conversion of psychosis in at risk individuals (Bodatsch et al., 2011; Perez et al., 2014b). In addition, MMN is sensitive to pharmacologic (Javitt et al., 1996; Umbricht et al., 2000, 2002; Engeland et al., 2002; Inami et al., 2005, 2007; Baldeweg et al., 2006; Dunbar et al., 2007; Lavoie et al., 2007; Ehrlichman et al., 2008; Martin et al., 2009; Dulude et al., 2010; Nakamura et al., 2011; Gil-da-Costa et al., 2013; Preskorn et al., 2014) and cognitive challenges (Rissling et al., 2013b) and predicts a positive response to some treatments (e.g., Kawakubo et al., 2007).

This paper demonstrates the feasibility of using MMN (and P3a) in multi-site studies of SZ patients and HCS and multi-site treatment trials with centralized quality assurance, data management, and analysis. The COGS-2 platform provided an unprecedented volume of subjects that allowed for the disentangling and discovery of participant-related characteristics that account for significant proportions of variability in MMN and P3a neurophysiologic biomarkers. We found that a relatively simple, 2-channel system yielded 91% usable MMN and P3a data in less than 30 min. The observed results replicate findings of highly significant MMN ($d = 0.96$) and P3a ($d = 0.93$) amplitude reductions in SZ comparable to those previously obtained in a single-laboratory with no appreciable differences across sites, further validating the use of these measures in multi-site studies. The large sample size also allowed for the identification and characterization of demographic factors that significantly contribute to MMN and P3a amplitudes. Significant clinical, cognitive, and functional correlates were observed. These results have implications for future genomic and clinical outcome studies.

It is important to emphasize that since ERP measures were added to the established COGS-2 study, many of the commonly used optimal environmental characteristics used in EEG studies could not be imposed on the sites (e.g., sound proofed and/or electrically shielded rooms, on-site expertise in electrophysiology). In fact, research assistants received a cumulative total of only 2.5 h of face-to-face training annually for both MMN/P3a and P3b paradigms (P3b findings reported separately in this issue by Turetsky et al.), though this training was embedded in the context of comprehensive 2-day assessment workshops. This 2.5 h training encompassed general EEG assessment principles, electrode

![Fig. 7](image1.png)

**Fig. 7.** Mean mismatch negativity (MMN) and P3a amplitude plotted Healthy Comparison Subjects (HCS) and schizophrenia (SZ) patients by use of anticholinergic medication. MMN and P3a values are corrected for age, sex, and race. Errors bars indicate 1 SE.

![Fig. 8](image2.png)

**Fig. 8.** Mean mismatch negativity (MMN) and P3a amplitude plotted Healthy Comparison Subjects (HCS) and schizophrenia (SZ) patients by smoking status. MMN and P3a values are corrected for age, sex, and race. Errors bars indicate 1 SE.

<table>
<thead>
<tr>
<th></th>
<th>MMN</th>
<th>P3a</th>
</tr>
</thead>
<tbody>
<tr>
<td>P3a</td>
<td>$-0.28^{	ext{***}}$</td>
<td></td>
</tr>
<tr>
<td>MMSE</td>
<td>$-0.14^{	ext{**}}$</td>
<td>0.05</td>
</tr>
<tr>
<td>SAPS</td>
<td>0.08</td>
<td>$-0.02^*$</td>
</tr>
<tr>
<td>SANS</td>
<td>0.10</td>
<td>0.02</td>
</tr>
<tr>
<td>Age of onset</td>
<td>$-0.09^*$</td>
<td>0.07*</td>
</tr>
<tr>
<td>Scale of Functioning</td>
<td>$-0.09^*$</td>
<td>0.07*</td>
</tr>
<tr>
<td>UPSA-B</td>
<td>$-0.16^{	ext{**}}$</td>
<td>0.06</td>
</tr>
<tr>
<td>Role functioning</td>
<td>$-0.13^{	ext{***}}$</td>
<td>0.07</td>
</tr>
</tbody>
</table>

**Table 2**

Correlations between demographically adjusted mismatch negativity (MMN) and P3a amplitudes with clinical, cognitive, and functional measures.

---

$^{	ext{***}} = p < 0.001$.  
$^{	ext{**}} = p < 0.01$.  
$^* = p < 0.05$. 

This research was supported by National Institute of Mental Health (NIMH) grants 6R01MH084706, 6R01MH086691, 6R01MH094552, 6R01MH096421, 6R01MH098847, 6R01MH098112, 6R01MH101234, 6R01MH102091, and 6R01MH103819.
application, impedance measurement and reduction, equipment installation and configuration, software setup and operation, data quality monitoring and patient instruction, and data upload procedures. Despite these apparent limitations, usable data was obtained from 91% of tested participants—an encouraging finding that should accelerate efforts to incorporate EEG biomarker testing in “real-world” clinical settings. This high rate of usable data could likely be further improved with additional training and oversight to correct some avoidable problems that occurred with test administration.

The expected patterns of results replicate previous findings from studies conducted in single-site, EEG specialty laboratories that benefit from onsite technical expertise, sophisticated high-density recording equipment, and environmental controls (e.g., dedicated sound-isolated recording suites). First, SZ patients exhibited uncorrected effect size deficits equivalent in magnitude to our previously published studies that used identical stimulation parameters, but different equipment (Light and Braff, 2005a; Light et al., 2007, 2012; Kang et al., 2009; Jahshan et al., 2012; Rissling et al., 2012; Takahashi et al., 2012) and meta-analyses (Umbricht and Krljes, 2005). Deficits of this large magnitude are visually evident in both group averages as well as in individual subject data obtained from all 1630 patients as shown in Fig. 2. Second, demographic factors – especially age – are strong contributors to MMN and P3a. Indeed, these factors collectively account for 26 and 18% of the variance in MMN and P3a, respectively, and also extend our previous findings using this paradigm (Kiang et al., 2009; Rissling et al., 2012; Takahashi et al., 2012). Third, sex and race also appear to contribute to MMN and P3a independent of the expected age and group effects. Specifically, females exhibited slightly larger ERP amplitudes. Likewise, race appears to be a significant predictor of MMN and P3a; slightly smaller amplitudes were detected in African Americans. Notably, impedance values were also significantly higher in African Americans which may have resulted in reduced signal-to-noise ratios of EEG, thereby attenuating ERP amplitudes. Fourth, it is important to note that several significant correlations were detected with demographically adjusted ERP amplitudes with the clinical, cognitive, and functional characteristics of the SZ patients; these correlations replicate and extend previous findings and further emphasize the utility of this approach. Specifically, smaller MMN amplitude was associated with an earlier age of illness onset, more severe positive and negative symptoms, worse global cognitive function, reduced functional capacity, functional status, and role functioning. Similarly, reduced P3a was significantly associated with an earlier age of illness onset and worse psychosocial functional status.

Demographic adjustment is a standard practice in clinical neuropsychological assessments of brain function (e.g., Heaton et al., 1991; Norman et al., 2011) and a similar strategy will clearly be necessary should these promising ERP biomarkers be used to guide assignment of patients to clinical interventions (Light and Näätänen, 2013; Light and Swerdlow, 2014; Perez et al., 2014a). A dramatic example of the importance of the demographic corrections can be drawn from the expanding number of studies demonstrating that MMN can improve the identification of children at highest imminent risk for developing psychosis (Nagai et al., 2013a). This literature raises the tantalizing possibility that a simple laboratory assay, alone or in conjunction with other measures, can contribute to prophylactic or early intervention strategies. While this hope appears reasonable, the risk of misclassification by failing to account for demographic factors can have profound deleterious effects.

Specific design considerations can be employed to minimize the impact of potential demographic confounds. For example, in repeated measurements designs an individual serves as his or her own control across time or conditions. Such designs would allow for investigators (or clinicians) to identify individuals who experience a significant change in ERP amplitudes over time. Thus, children being monitored due to being at elevated risk for developing a psychotic disorder may exhibit a significant change in MMN amplitudes suggestive of a decline in brain function. Conversely, patients who exhibit malleability of ERPs in response to targeted cognitive and/or pharmacologic interventions could be selected for further treatments (Light and Näätänen, 2013; Light and Swerdlow, 2014; Perez et al., 2014a). If repeated measures designs are suboptimal for a specific application, the use of homogeneous subgroups carefully matched in at least age – the most robust demographic covariate of MMN and P3a components – should be considered.

There are caveats and cautionary notes for interpreting this extensive characterization of MMN and P3a in SZ patients and HCS. First, the analyses presented in this manuscript represent an initial examination in a rich and fertile database. There are likely other factors (e.g., substance use histories) that also contribute to variability in MMN and P3a amplitudes that require examination in future analyses. Second, the extensive COGS-2 assessment was cross-sectional and does not lend itself to identifying cause–effect relationships. For example, does the finding that unmedicated SZ patients have larger MMN and P3a amplitudes suggest that antipsychotics “cause” a reduction in these measures or do patients who do not take or perhaps even require antipsychotics (and are presumably less symptomatic and higher functioning) have larger (more “normal”) amplitudes? Post-hoc analyses revealed that these antipsychotic medication subgroups also significantly differed on other key clinical, cognitive, and functional domains – an issue for future planned analyses. Third, some statistically significant findings yielded small effects (e.g., smoking and anticholinergic medications) or relied upon imprecise (perhaps even anachronistic) subgrouping such as self-identified race; care should be taken to not over-interpret such unreplicated findings. Fourth, the use of only 2 channels in this simplified system, while highly practical for non-specialty centers, obviously precludes the ability to perform EEG analyses that require higher-density sensor arrays such as examination of the neural substrates in clinical populations (Takahashi et al., 2012; Rissling et al., in press).

In conclusion, this study demonstrates that MMN and P3a ERP biomarkers can be feasibly added to multi-site clinical studies without the usual constraints for most high-density EEG studies conducted in academic EEG specialty centers. The absence of site differences supports the use and further development of ERP applications for use in real-world community care centers. As with many clinical tests of brain function, demographic factors, particularly age, contribute to MMN and P3a amplitudes and should be carefully considered in future biomarker-informed clinical studies (Light and Näätänen, 2013; Light and Swerdlow, 2014; Perez et al., 2014a). Demographic corrected MMN and P3a amplitudes exhibit significant correlations with well-established measures of clinical, cognitive, and psychosocial functioning in SZ patients, which replicate and extend previous studies and further underscore the utility of these measures. This platform will serve as a valuable resource for future functional and genomic analyses.

Role of funding source

Other than providing support, the National Institute of Health does not have any further role in this manuscript.

Contributors

Dr. Light provided training and ongoing quality assurance for MMN and P3a measures acquired at all COGS-2 sites. Dr. Light also processed and analyzed all EEG data and wrote the manuscript. Drs. Thomas, Lazzeroni, Sugar, and Light contributed to the statistical analysis plan. All other authors participated in aspects of study design, including subject recruitment, EEG testing, and validation of the clinical and endophenotype data. All authors were responsible for reviewing, editing, and approving the final version of the manuscript.

Conflicts of interest

Dr. Light reports having been a consultant to EnVivoForum and Astellas and serves on an advisory board for Neuroverse. Dr. Green has been a consultant to AbbVie, Biogen, DSP, EnVivoForum and Roche, and he is on the scientific advisory board of Mnemosyne. He has received research funds from Amgen. Dr. Lazzeroni is an inventor on a patent application filed by Stanford University on genetic polymorphisms associated with depression. Dr. Nuechterlein has received unrelated research support from Janssen Scientific Affairs, Genentech, and Brain Plasticity, Inc., and has consulted to Genentech, Otsuka, Janssen, and...
and Brain Plasticity, Inc. Dr. Swerdlow has been a consultant for Geno Sciences, Ltd. All other authors declare that they have no conflict of interest.

Acknowledgments

The authors wish to thank all of the participants and support staff that made this study possible, including the following key personnel:

University of California San Diego (R01-MH065571; MH042228, MH079777, MH087889, Brain Behavioral Research Foundation, Sidney R. Baer Jr. Foundation): Joyce Sprock, Barbara Haugland, Lauren Belleville, Stacy Langton, Daniel Mathias, Natalie McCarthy, Marlena Pela, Erich Riesen, and Maria Bongiovanni

Mount Sinai School of Medicine (R01-MH065554): Rui Ferreira, Carolyn Khanian, Denise Poche-Jetter, and Rebecca West

University of California Los Angeles (R01-MH67507): William Horam, Mark Sergi, Amanda Binder, Lusine Garapetian, Robert Hubert, Heidi Kupping, Thrin Luu, Ian Mathis, Mark McGee, Anacelli Myers, Felice Reddy, Amber Twidell, Christen Wadson, and Katie Weiner

University of Pennsylvania (R01-MH65578): Kate Cassidy, Erich Dress, Colin Gallagher, Mary March, Kathleen McKenna, Alison Mott, Michael Pato, Jan Richards, Kosha Ruparel, and Chandni Singh

University of Wisconsin (R01-MH65558): Amy Cassidy, Erich Dress, Colin Gallagher, Mary March, Kathleen McKenna, Alison Mott, Michael Pato, Jan Richards, Kosha Ruparel, and Chandni Singh

References


