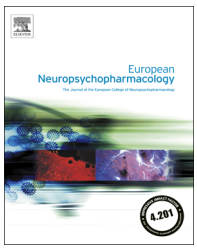
European Neuropsychopharmacology (2015) 25, 1981–1990

www.elsevier.com/locate/euroneuro



Utility of event-related potentials in predicting antidepressant treatment response: An iSPOT-D report

Rik van Dinteren[a,b], Martijn Arns[b,c,*], Leon Kenemans[c], Marijtje L.A. Jongsma[d], Roy P.C. Kessels[a], Paul Fitzgerald[e], Kamran Fallahpour[f,g], Charles Debattista[h], Evian Gordon[i], Leanne M. Williams[h,j]

[a]*Donders Institute for Brain Cognition and Behaviour, Radboud University Nijmegen, Nijmegen, The Netherlands*
[b]*Research Institute Brainclinics, Nijmegen, The Netherlands*
[c]*Department of Experimental Psychology, Utrecht University, Utrecht, The Netherlands*
[d]*Behavioural Science Institute, Radboud University Nijmegen, Nijmegen, The Netherlands*
[e]*Monash Alfred Psychiatry Research Centre, Monash University Central Clinical School and the Alfred, Melbourne, Vic., Australia*
[f]*Department of Psychiatry at the Icahn School of Medicine at Mount Sinai, New York, NY, USA*
[g]*Brain Resource Center, New York, USA*
[h]*Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, CA, USA*
[i]*Brain Resource, Sydney, NSW, Australia and San Francisco, CA, USA*
[j]*Veterans Affairs Palo Alto Healthcare System, and the Sierra Pacific Mental Illness, Research, Education, and Clinical Center (MIRECC), Palo Alto, CA, USA*

Received 12 April 2015; received in revised form 3 July 2015; accepted 28 July 2015

KEYWORDS
Event-Related Potential (ERP);
Major Depressive Disorder (MDD);
Antidepressants;
Treatment prediction

Abstract
It is essential to improve antidepressant treatment of major depressive disorder (MDD) and one way this could be achieved is by reducing the number of treatment steps by employing biomarkers that can predict treatment outcome. This study investigated differences between MDD patients and healthy controls in the P3 and N1 component from the event-related potential (ERP) generated in a standard two-tone oddball paradigm. Furthermore, the P3 and N1 are investigated as predictors for treatment outcome to three different antidepressants.
In the international Study to Predict Optimized Treatment in Depression (iSPOT-D) – a multicenter, international, randomized, prospective practical trial – 1008 MDD participants were

*Correspondence to: Research Institute, Brainclinics Bijleveldsingel 34, 6524 AD Nijmegen, The Netherlands. Tel.: +31 24 7503505; fax: +31 24 8901447.
E-mail address: martijn@brainclinics.com (M. Arns).

http://dx.doi.org/10.1016/j.euroneuro.2015.07.022

randomized to escitalopram, sertraline or venlafaxine-XR. The study also recruited 336 healthy controls. Treatment response and remission were established after eight weeks using the 17-item Hamilton Rating Scale for Depression. P3 and N1 latencies and amplitudes were analyzed using a peak-picking approach and further replicated by using exact low resolution tomography (eLORETA).

A reduced P3 was found in MDD patients compared to controls by a peak-picking analysis. This was validated in a temporal global field power analysis. Source density analysis revealed that the difference in cortical activity originated from the posterior cingulate and parahippocampal gyrus. Male non-responders to venlafaxine-XR had significantly smaller N1 amplitudes than responders. This was demonstrated by both analytical methods. Male non-responders to venlafaxine-XR had less activity originating from the left insular cortex.

The observed results are discussed from a neural network viewpoint.

1. Introduction

The World Health Organization (WHO) has ranked major depressive disorder (MDD) as the leading cause of disability worldwide (Marcus et al., 2012). Individual responses to antidepressants vary which often leads to the need to try different types of medication using a 'trial-and-error' approach. The STAR*D trial (Fava et al., 2003; Rush et al., 2006) demonstrated that this stepwise approach left one third of patients resistant to treatment after being randomized to three different treatments, also with different modes of action. This is not sufficiently effective and it is essential to improve antidepressant treatment response in MDD. One way this could be achieved is by reducing the number of treatment steps and guiding the right patients to the right treatment using biomarkers. This is referred to as 'personalized medicine' or 'precision medicine'.

1.1. Treatment prediction employing event-related potentials

A potential predictor for treatment outcome lies in MDD patients' electrophysiological properties. The event-related potential (ERP) is a waveform in brain activity that is time-locked to an event, for instance an auditory stimulus. Some of the ERPs' components have been studied for their predictive value in antidepressant treatment response, where most research has been focused on the P300 (P3) and N100 (N1) (for review see: Arns and Olbrich, 2014; Olbrich and Arns, 2013).

The P3 consists of neural activity originating from several origins, including presumably the prefrontal cortex, the temporoparietal junction, the primary auditory cortex and possibly other sources (Friedman, 2003). The P3 can be measured across the scalp, but he P3 measured at a frontal electrode site does not necessarily reflect the same cortical activation as the P3 measured at a posterior site. For older participants compensatory activation from frontal regions may be involved which may be represented by the frontal P3 and not the parietal P3 (van Dinteren et al., 2014a).

Most studies, but not all, have found lower P3 amplitudes in MDD patients compared to healthy controls. Some studies reported only non-significant trends of reduced P3 amplitudes in MDD patients (Bruder et al., 2009).

So far, research on antidepressant treatment response involving the P3 has been ambiguous. There are studies that found responders to antidepressants having larger P3 amplitudes than non-responders (Bruder et al., 1995; Jaworska et al., 2013). In contrast, responders to treatment with repetitive Transcranial Magnetic Stimulation (rTMS) were found to have lower P3 amplitudes than non-responders, although only marginally significant (Arns et al., 2012). Regarding P3 latency, the results have been mixed as well. Some found no effect of P3 latency (Jaworska et al., 2013), while other studies found slower P3's in non-responders to antidepressants (Işıntaş et al., 2012; Kalayam and Alexopoulos, 1999; Vandoolaeghe et al., 1998).

Spronk and colleagues found that a larger N1 amplitude at baseline (assessed in an oddball paradigm) was associated with better antidepressant treatment outcome (Spronk et al., 2011), suggesting that maybe the oddball N1 amplitude could serve as a predictor to treatment outcome, possibly in line with earlier research implicating the N1 component of the Loudness Dependent Auditory Evoked Potential (LDAEP) in prediction of treatment outcome (see Kenemans and Kähkönen, 2011; Olbrich and Arns, 2013 for review).

In this study we used data from the multi-center, randomized, prospective open-label international Study to Predict Optimized Treatment Response in Depression (iSPOT-D) (see Williams et al. for details). The aim of this study was to investigate differences between MDD patients and healthy controls in the P3 and N1 component from the ERP generated in a standard two-tone oddball paradigm. Furthermore, the P3 and N1 were investigated for their predictive power of antidepressant treatment response. In addition, we tested if P3 and N1 measures can aid in differential prediction to the three medication types. We expected to find (1) reduced P3 amplitudes in MDD patients, (2) larger P3 amplitudes in treatment responders compared to non-responders and (3) larger N1 amplitudes in treatment responders compared to non-responders. In our analysis we have first used a traditional peak-picking method to quantify the ERP components and also obtain latencies for all components. In addition, we have applied a time-based analysis using eLORETA to replicate the findings obtained from the peak-picking method and to also obtain localizations of the observed effects.

This study is among one of several planned iSPOT-D studies that were approved by the iSPOT-D publication committee. Other planned analyses on EEG metrics are

reported in separate manuscript such as for example Arns et al. (2015a, 2015b) in this journal, for an investigation of the role of frontal and rostral anterior cingulate theta EEG power in antidepressant treatment prediction.

2. Experimental procedures

2.1. Design

Data from this study come from the iSPOT-D study, which is an international multi-center, randomized, prospective open-label phase IV clinical trial. MDD patients were randomized to escitalopram, sertraline or venlafaxine-XR in a 1:1:1 ratio. The study was designed to approximate real-life treatment of MDD and therefore no placebo condition was included. A complete description of the study has been published earlier (Williams et al., 2011).

2.2. Participants and treatment

There were 1008 MDD patients and 336 healthy controls (HC) included in this study. All participants were aged 18–65 years. MDD patients were treatment-naïve or medication was washed-out (5 half-lives). At baseline, a primary diagnosis of unipolar, non-psychotic MDD was confirmed by administration of the Mini-International Neuropsychiatric Interview (Sheehan et al., 1998) according DSM-IV criteria and a score ≥ 16 on the Hamilton Rating Scale for Depression ($HRSD_{17}$). The visit consisted of questionnaires, an EEG assessment, a neuropsychological battery and a blood draw. After the baseline visit, MDD patients were randomized to one of the three antidepressants and a follow-up visit was scheduled after eight weeks. All baseline tests, except for the blood draw and the diagnostics, were repeated at the follow-up visit. Although this study only includes data from the baseline and week 8 visits, participants were in fact followed-up for one year in total.

This study was approved by the institutional or ethics review boards at all of the participating sites. This trial was registered with ClinicalTrials.gov. Registration Number: NCT00693849; URL: http://clinicaltrials.gov/ct2/show/NCT00693849.

2.3. Electroencephalographic data acquisition

EEG recordings were performed using a standardized methodology and platform (Brain Resource Ltd., Australia). Details of this procedure have been published elsewhere (Arns et al., 2008; Williams et al., 2011) and details of the reliability and across-site consistency of this EEG procedure have been published(Paul et al., 2007; Williams et al., 2005). In summary, participants were seated in a sound and light attenuated room that was controlled at an ambient temperature of 22 °C. EEG data were acquired from 26 channels: Fp1, Fp2, F7, F3, Fz, F4, F8, FC3, FCz, FC4, T3, C3, Cz, C4, T4, CP3, CPz, CP4, T5, P3, Pz, P4, T6, O1, Oz and O2 (Quikcap with sintered Ag/AgCl electrodes; Neuroscan NuAmps DC amplifier; extended 10–20 electrode international system). Data were offline re-referenced to averaged mastoids with a ground at AFz. Horizontal eye movements were recorded with electrodes placed 1.5 cm lateral to the outer canthus of each eye. Vertical eye movements were recorded with electrodes placed 3 mm above the middle of the left eyebrow and 1.5 cm below the middle of the left bottom eyelid. Skin resistance was aimed at <5 kOhms for all electrodes. A continuous acquisition system was employed and EEG data were electrooculogram (EOG)-corrected offline (Gratton et al., 1983). The sampling rate of all channels was 500 Hz. A low pass filter with an attenuation of 40 dB per decade above 100 Hz was employed prior to digitization.

The oddball paradigm consisted of a quasi-random sequence of 280 frequent background tones (500 Hz) and 60 infrequent target (1000 Hz) tones. Two targets could not appear consecutively. All stimuli (50 ms; 5 ms rise and fall time) were presented binaurally at a volume of 75 dB SPL with an inter-stimulus interval of 1000 ms. Participants were instructed to press two buttons simultaneously (one for each index finger to counterbalance motor effects) when they heard a target tone and to ignore the background tones. Speed and accuracy of response were both equally stressed in the instructions. Before the actual test they were presented with a brief practice run to clarify the distinction between the two tones.

2.4. ERP scoring

Conventional ERP averages were generated at all electrode sites for target stimuli. Only stimuli with a correct response were included in the average. Before averaging each single trial ERP was filtered with a Tukey (tapered cosine) window. No signal was passed above 35 Hz. For the target stimuli waveforms, the peaks (amplitude and latency) of the N1, P2, N2 and P3 ERP components were identified, relative to a 300 ms pre-stimulus baseline, at Fz and Pz (to capture possible compensatory effects). Using pre-determined latency windows, peaks were semi-automatically scored by BrainVision Analyzer 2: N1 (70–120 ms) and P3b (220–550 ms) and manually corrected by the first author (RvD). All scoring was checked by a second rater, and disagreements in scoring ($<5\%$) were settled by the second author (MA).

2.5. Current source density analysis

eLORETA software and statistical nonparametric mapping method (SnPM) (Pascual-Marqui, 1999, 2002, 2007; Pascual-Marqui et al., 1994) were used (1) to replicate obtained results from the ERP peak-picking analysis by using a time-based analysis, and (2) to localize differences in brain activity between groups to cortical sources. This was done by analyzing time frames of averaged group ERPs with a *t*-test (Nichols and Holmes, 2002). Those timeframes that significantly differ in global field power were entered in a *t*-test on log-transformed current source density data in order to obtain the cortical sources responsible for the group difference in brain activity (Strik et al., 1998). eLORETA has been validated in multiple studies and the standard 10/20 EEG system employing electrode montages with for example 25 electrodes has been proven sufficient for source localization (Pascual-Marqui et al., 2002). Statistics from both methods are corrected for multiple comparisons by means of a randomization procedure.

2.6. Statistics

Response was defined as a $>50\%$ decrease in $HRSD_{17}$ score from baseline to week 8. In this analysis, we primarily assessed responders vs. non-responders. Differences in age and sex were tested using a *t*-test or non-parametric tests (sex). In case of group differences in one of these measures, these variables were added as a covariate.

For comparison of MDD vs. healthy controls as well as investigating treatment prediction a repeated measures ANOVA was conducted for P3 and N1 amplitude and latencies with within-subject factors Site (Fz or Pz), between-subject factors group (MDD vs. Controls or Response vs. Non-Response), treatment arm (For response analyses only: ESC, SER and VEN), Age (Young vs. Old) and sex. When significant interactions were found, univariate analyses were performed. The cut-off for young and old participants was set at 46 years based on age-related development of P3 amplitude as reported in our previous studies (van Dinteren et al., 2014a, 2014b).

A partial correlation (correcting for age) was run between the percentage improvement on the $HRSD_{17}$ between baseline and week 8, $HRSD_{17}$ at intake and $HRSD_{17}$ at week 8 and the significant ERP components obtained.

Table 1 Demographic features of MDD patients and controls and treatment outcomes for patients who completed treatment per protocol.

	Whole sample		Per protocol completers			Early drop-outs		
	MDD	*Controls*	*Escitalopram*	*Sertraline*	*Venlafaxine-XR*	*Escitalopram*	*Sertraline*	*Venlafaxine-XR*
Number	1008	336	217	234	204	119	102	132
Females	571	191	119	139	120	59	52	82
Average Age (years)	37.84 (12.57)	36.99 (13.08)	38.85 (12.47)	38.34 (12.42)	38.46 (12.87)	37.28 (12.93)	36.77 (11.85)	35.68 (12.65)
$HRSD_{17}$ Baseline	21.88 (4.12)	1.15 (1.63)	21.75 (3.98)	21.95 (4.17)	21.50 (3.89)	21.77 (4.27)	21.89 (4.18)	22.66 (4.41)
$HRSD_{17}$ Week 8	9.67 (6.43)	1.06 (1.43)	9.29 (6.60)	9.41 (6.08)	9.71 (6.22)	-	-	-
$HRSD_{17}$ Anxiety[a]	6.16 (1.99)	.57 (.97)	6.18 (1.95)	6.27 (1.96)	6.14 (1.85)	6.04 (2.19)	5.86 (2.19)	6.08 (2.11)
% Remission ($HRSD_{17}$)			48	47	44	-	-	-
% Response ($HRSD_{17}$)			60	67	63	-	-	-

Abbreviations: $HRSD_{17}$, 17-item Hamilton rating scale for depression; MDD, Major depressive disorder; XR, Extended release.
[a]Measured at baseline.

All statistics for treatment prediction were performed on data from MDD participants who completed 8 weeks of treatment per protocol: participants who were dosed with their randomized medication for a minimum of 6 weeks and who returned for their week 8 visit and were still receiving their randomized medication at this visit ('per protocol' grouping). Significance level was set at $p \leq .05$ and effect sizes (ES) of main effects are reported in Cohen's *d*.

3. Results

For the comparison between MDD and healthy controls, 1008 MDD patients were compared to 336 HC. For treatment prediction the 'per protocol' grouping consisted of 655 MDD participants. See Table 1 for demographic information and remission and response rates for the whole MDD group and the separate treatment groups used for treatment prediction. The overall remission and response rates in this MDD sample were respectively 46% and 64%. Treatment was randomized as follows: Escitalopram ($N=217$), Sertraline ($N=234$) and Venlafaxine-XR ($N=204$). The treatment groups did not differ regarding age, sex, baseline MDD severity ($HRSD_{17}$), anxiety severity ($HRSD_{17}$), remission or response rates.

3.1. MDD participants vs. controls

There was no significant difference regarding age between MDD participants ($M=37.84$, $SD=12.57$) and HC ($M=36.99$, $SD=13.08$); $t(1342)=-1.06$, $p=.29$. Also, there were no significant differences regarding sex between the two groups ($\chi^2(1,1344)=.004$, $p=.949$).

P3 amplitude
Repeated-measures ANOVA, yielded the following results for P3 amplitude: An effect of Site ($F(1,1087)=695.37$, $p<.001$), Sex ($F(1,1087)=29.06$, $p<.001$), a Group X Site interaction ($F(1,1087)=6.19$, $p=.013$) and a main effect of Group ($F(1,1087)=14.61$, $p<.001$).

The significant Group *x* Site interaction was further analyzed. Separate univariate analyses for Fz and Pz demonstrated a main effect for Group (Fz: $F(1,1091)=4.74$, $p=.03$; Pz: $F(1,1104)=18.14$, $p<.001$). MDD participants had significantly lower P3 amplitudes compared to HC at Fz ($M=3.90\pm5.48$ vs. $M=4.76\pm6.41$; $d=.14$) and Pz ($M=9.51\pm5.93$ vs. $M=11.73\pm7.31$; $d=.33$).

A significant correlation between P3 amplitude at Pz and depression severity at baseline ($HRSD_{17}$) was obtained within the group of MDD patients ($r=-.069$, df=803, $p=.024$). There were no significant correlations between P3 amplitude at Fz and baseline $HRSD_{17}$ in MDD patients nor a correlation for P300 at Pz and Fz and baseline $HRSD_{17}$ in healthy controls.

P3 latency
Repeated-measures ANOVA, yielded the following results for P3 latency: An effect of Site ($F(1,1087)=98.420$, $p<.001$), Sex ($F(1,1087)=16.21$, $p<.001$), Age ($F(1,1087)=39.592$, $p<.001$), a Group × Sex × Age interaction ($F(1,1087)=5.487$, $p=.019$) and no effect of Group ($p=.167$).

The interaction effect of Group × Sex × Age was further investigated. Separate repeated measures analyses revealed that only young females with MDD had slower P3 latencies than young female healthy controls ($F(1,433)=8.70$, $p=.003$, $d_{Fz}=-.29$, $d_{Pz}=-.28$). There were no effects of Group in older females, or young and old males.

N1 amplitude
Repeated-measures ANOVA, yielded the following results for N1 amplitude: An effect of Site ($F(1,1115)=551.479$, $p<.001$), Sex ($F(1,1115)=11.499$, $p<.001$), Age ($F(1,1115)=6.198$, $p=.013$), a Group × Sex × Age × Site interaction ($F(1,1115)=4.835$, $p=.028$) and an effect of Group ($F(1,1115)=11.423$, $p<.001$).

The Group × Sex × Age × Site interaction effect was further investigated by univariate analyses for both electrode sites per sub-group. The Group effect was found to be limited to young women on both electrode sites (Fz: *F*

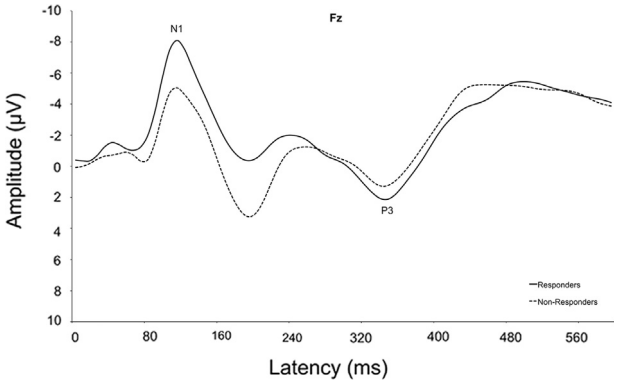


Figure 1 Averaged ERP measured at Fz of male responders and non-responders to venlafaxine-XR.

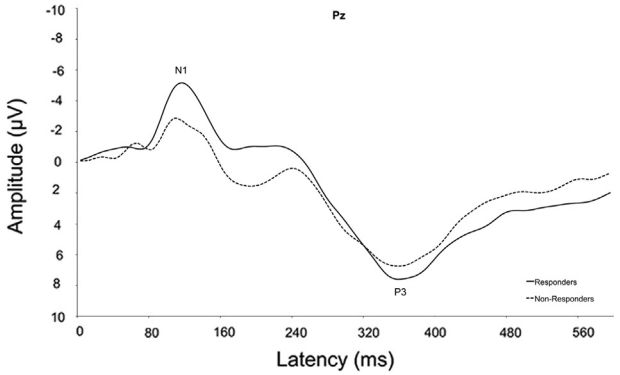


Figure 2 Averaged ERP measured at Pz of male responders and non-responders to venlafaxine-XR.

(1,444)=17.08, $p<.001$, $d=-.42$; Pz: $F(1,444)=14.16$, $p<.001$, $d=-.40$) and young men (only on Pz: $F(1,335)=5.56$, $p=.019$, $d=-.29$), where MDD patients had smaller N1 amplitudes as compared to HC.

N1 latency

Repeated-measures ANOVA, yielded the following results for N1 latency: An effect of Site ($F(1,1115)=5.993$, $p=.015$), Age ($F(1,1115)=7.883$, $p=.005$) and no effect of Group ($p=.300$).

3.2. Responders vs. non-responders ($HRSD_{17}$)

Responders did not differ from non-responders regarding sex, baseline MDD severity, baseline anxiety severity or type of treatment. However, responders were significantly younger than non-responders ($M=37.42\pm12.20$ vs. $M=40.51\pm12.95$; $t(653)=3.05$, $p=.002$). Since Age group was included as a between-subject factor, interactions with age group are included in the following analysis.

P3 amplitude and P3 latency

There were no effects nor interactions involving Response Type on the P3 amplitude (Group effect: $p=.980$) and P3 latency (Group effect: $p=.296$), suggesting no role of these measures in predicting antidepressant response.

N1 amplitude

Repeated-measures ANOVA, yielded the following results for N1 amplitude: An effect of Site ($F(1,527)=272.271$, $p<.001$), Sex ($F(1,527)=4.747$, $p=.030$), a Response Type × Sex interaction ($F(1,527)=5.750$, $p=.017$), a Response Type × Sex × Treatment arm interaction ($F(2,527)=3.052$, $p=.048$) and a trend effect for Response Type ($p=.054$).

The Response Type × Sex × Treatment Arm interaction effect was further investigated by separate RM ANOVAs by sex.

In males, main effects of Response Type ($F(1,222)=8.94$, $p=.003$, $d_{Fz}=.41$, $d_{Pz}=.24$), and Site were obtained. Furthermore, there were interaction effects of Response Type × Treatment Arm ($F(2,222)=5.43$, $p=.005$) and Response Type × Site ($F(1,222)=4.012$, $p=.046$) in males. Separate RM analyses per treatment arm revealed no effects of Response Type with escitalopram or sertraline. However, with venlafaxine-XR there was an effect of Response Type with large effect sizes ($F(1,68)=14.45$, $p<.001$, $d_{Fz}=.89$, $d_{Pz}=.73$), Site ($F(1,68)=49.01$, $p<.001$) and an interaction effect of Response Type × Site ($F(1,68)=5.04$, $p=.028$). Separating this analysis into two univariate analyses for each site demonstrated a main effect of Response Type on both sites (Fz: $F(1,71)=14.02$, $p<.001$, $d=.89$; Pz: $F(1,71)=11.12$, $p=.01$, $d=.73$). Male responders to venlafaxine-XR have larger N1 amplitudes than non-responders, see Figures 1 and 2.

In females, a main effect of Site ($F(1,305)=143.38$, $p<.001$) and an interaction effect of Site × Age × Treatment Arm × Response Type ($F(2,305)=3.75$, $p=.025$) were obtained. In all possible subgroups (young/old using escitalopram/sertraline/venlafaxine-XR) there was no effect of Response Type on the frontal or parietal N1 amplitude in females.

Within this group of males treated with venlafaxine-XR, N1 amplitude did not correlate with baseline $HRSD_{17}$ for both Fz and Pz ($p>.770$), however N1 amplitude at Fz did correlate significantly with $HRSD_{17}$ at week 8 ($p=.043$; $r=.241$; DF=69) and percentage improvement on $HRSD_{17}$ ($p=.030$; $r=.258$; DF=69).

Splitting the Response Type × Treatment Arm interaction into separate RM ANOVAs per treatment arm ultimately led to the same conclusions.

N1 latency

Repeated-measures ANOVA, yielded the following results for N1 latency: An effect of Site ($F(1,527)=14.981$, $p<.001$) and a Response Type × Age group × Sex × Treatment arm interaction ($F(2,527)=3.358$, $p=.036$) and no effect for Response Type ($p=.835$).

The Response Type × Age group × Sex × Treatment arm interaction effect was further investigated. Of all possible sub-groups (young/old males/females using escitalopram/sertraline/venlafaxine-XR) there was no significant effect of Response Type.

Post-hoc analysis of behavioral data

In a recent study on the same iSPOT-D study, cognitively impaired patients were found to have poorer treatment outcome (Etkin et al., 2014). Therefore, post-hoc analyses including the cognitive impairment grouping factor (intact/impaired) from that study data were run. Inclusion of cognitive impairment as a factor did not alter the earlier results. In short, cognitive impairment did not refine earlier obtained results.

Table 2 Results of a logistic regression model with N1 amplitude at Fz as a predictor for Response Type.

			95% confidence interval for odds ratio		
	B(SE)	*p*	*Lower*	*Odds Ratio*	*Upper*
Constant	−1.493				
N1 amplitude (Fz)	−.275 (.087)	.002	.640	.759	.901

Note: R^2=.14 (Hosmer and Lemeshow), .18 (Cox and Snell), .24 (Nagelkerke). Model $\chi^2(1)=13.95$, $p<.001$.

Post-hoc logistic regression analysis
In a subgroup of males that used venlafaxine-XR, a logistic regression analysis was performed with Response Type as dependent variable. Age (continuous), N1 amplitude (Fz) and N1 amplitude (Pz) were simultaneously entered as predictors. Age (continuous) and N1 amplitude (Pz) did not contribute significantly to the model and were removed. The logistic regression model with one predictor, N1 amplitude (Fz) predicted 68% of Response Type. There were 58% of the non-responders and 76% of the responders predicted correctly. The odds ratio for N1 amplitude (Fz) indicates that smaller values (i.e. *higher* negative amplitudes) increase the probability of being a responder, see Table 2.

eLORETA analyses
The observed effects obtained with the peak-picking method were replicated using eLORETA analyses. In addition the eLORETA analyses provided cortical source localization of the observed P3 and N1 differences.

MDD vs HC
The eLORETA software was applied to ERP data of MDD patients and healthy controls to find time points in the EEG where they significantly differed ($t>3.878$, $p<.05$). This was located at 344–384 ms post-stimulus which corresponds to the P3 interval. Between-group comparisons of the averaged current density distributions for this time segment were performed using a *t*-test on the log-transformed current density values. MDD patients' P3s consist of less activity originating from the posterior cingulate and parahippocampal gyrus (BA31, $t>3.498$, $p<.05$, 2-tailed test), see Figure 3.

Temporal source analysis with eLORETA localized the difference regarding the N1 between MDD patients and healthy controls at 104–120 ms post-stimulus. However, between-group comparison of averaged current density distributions within this timeframe yielded no significant differences in source space.

Male responders to venlafaxine-XR vs. non-responders
Based on the peak-picking results, the same analysis was conducted for male responders vs. non-responders to venlafaxine-XR within the N1 interval (70–120 ms). The EEGs of male responders to venlafaxine-XR differed significantly from male non-responders to venlafaxine-XR at 88–104 ms post-stimulus ($t>3.38$, $p<.05$). Between-group comparisons of the averaged current density distributions of the time segment 88–104 ms were performed using a *t*-test on the log-transformed current density values. Male non-responders to venlafaxine-XR had significantly less activity originating from the left insula cortex, BA13 ($t=-3.72$, $p=.048$), see Figure 4.

4. Discussion

In this study differences between MDD patients' and healthy controls' P3 and N1 components of the auditory oddball ERP were investigated. Furthermore, these components were investigated for their value in predicting antidepressant treatment response. For validation purposes, two different analytic methods were used and both methods confirmed the main results. Additionally, CSD analyses were used to localize cortical sources of differences in brain activity between groups.

4.1. MDD and P3

The P3 amplitude was smaller in MDD patients compared to healthy controls, with small effect sizes ($d_{Fz}=.14$; $d_{Pz}=.33$), and within the MDD sample P3 amplitude at Pz also correlated with MDD severity. A temporal global field power analysis confirmed this finding by significant global field power differences between the groups in a time interval that corresponded to the P3. Using the largest sample size to date, this study confirms other findings of low P3 amplitudes in MDD patients (Gangadhar et al., 1993; Kawasaki et al., 2004; Kemp et al., 2009; Röschke and Wagner, 2003). In addition, we were able to localize the difference in P3 amplitude between the groups to specific cortical structures, namely the posterior cingulate and the parahippocampal gyrus (PHG). The posterior cingulate is a large posterior hub in the default mode network (DMN) that has consistently been found to be abnormal in MDD (Liston et al., 2014). In this view, the PHG serves as a central hub within a sub-network that links the DMN to the hippocampus and thereby facilitates in memory (Ward et al., 2014). Cognitive problems associated with MDD (attention, working memory) might be related to inefficient information processing originating from suboptimal connections between the PHG sub-network and the DMN. The P3 amplitude might be a biomarker for indexing the neural interaction between the DMN and the PHG sub-network. This is highly speculative and requires further replication, but it would impose a new interesting view on the P3.

For the N1 it was found that MDD patients had low N1 amplitude, and this was specifically found for young subjects (age <46 yr). This is in agreement with (Coffman et al., 1989), albeit other studies have not found differences in N1 amplitude (Bruder et al., 1995; Bruder et al., 1998; Kawasaki et al., 2004; Urretavizcaya et al., 2003). The effect was validated by significant differences in global field power between the groups at a time point that corresponds to the time window of the N1, however, no significant differences in cortical current source densities between young MDD patients and young healthy controls at this N1 time window were found.

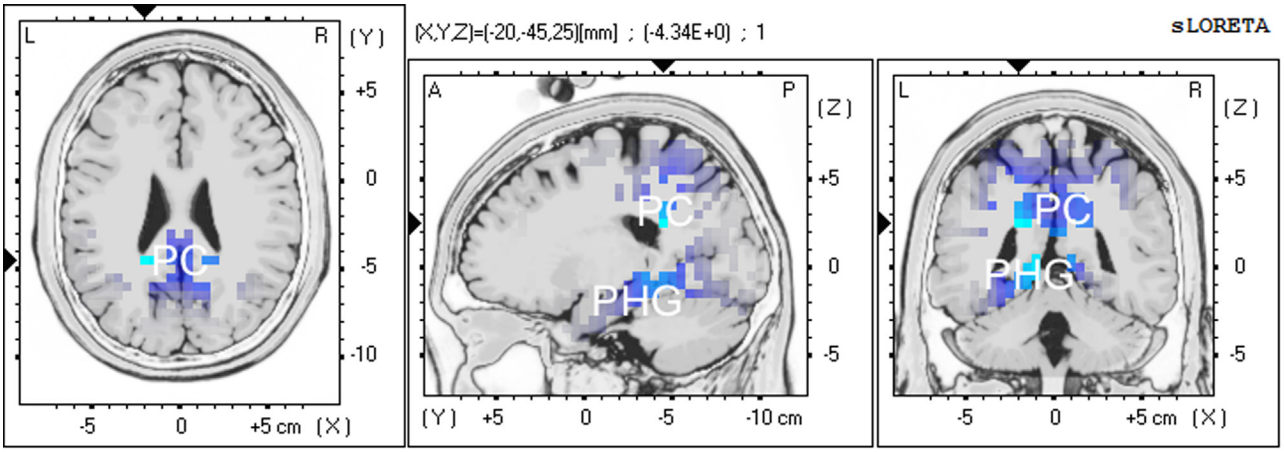


Figure 3 MDD patients' P3's consist of significantly less activity originating from the posterior cingulate (PC) and parahippocampal gyrus (PHG).

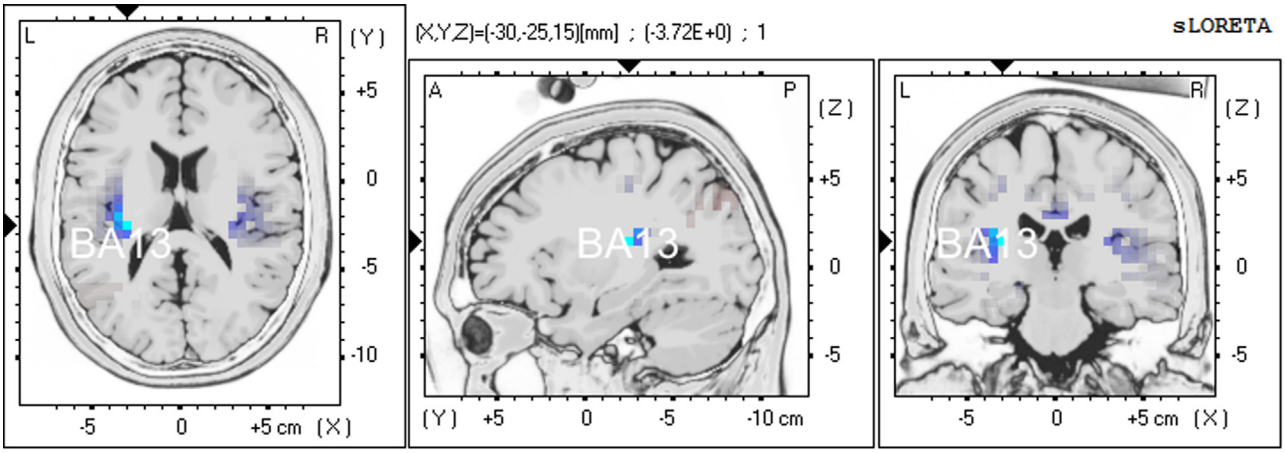


Figure 4 Male non-responders to venlafaxine-XR have significantly less activity originating from the left insular cortex (BA13) at 88-104 ms post-stimulus (the N1 time frame).

4.2. Treatment prediction

For antidepressant response prediction, we could not replicate earlier reports of differences in P3 amplitude between responders and non-responders to antidepressant treatment (Bruder et al., 1995; Jaworska et al., 2013). Although, it has to be noted that these earlier reports used paradigms that were cognitively more challenging that the standard oddball paradigm from this study. Such paradigms might be more sensitive to predicting treatment response. Furthermore, the treatments that were investigated in these studies differed from our study. However, we did find male responders to venlafaxine-XR demonstrating larger (more negative) pre-treatment N1 amplitudes than non-responders, with large effect sizes (d_{Fz}=.89, d_{Pz}=.73) and N1 amplitudes at Fz correlated significantly with percentage improvement on $HRSD_{17}$. Furthermore, this finding replicates our earlier study in 25 MDD patients using identical methods and the exact same oddball paradigm where this same effect was found for the whole group, including males and females (Spronk et al., 2011). The source of this difference in N1 amplitude was localized to the insular cortex.

The insula has been associated with MDD and with antidepressant treatment response in other studies. In PET studies, hypometabolism of the anterior insula, among other regions, was associated with response to fluoxetine (Mayberg et al., 2002), while increased metabolism of the insula was associated to remission after treatment with escitalopram (McGrath, 2013). Using fMRI, increasing activation of the insula as venlafaxine treatment progressed was observed in MDD patients (Davidson et al., 2003). In another fMRI study where unmedicated MDD patients performed a task requiring attention to bodily senses, insula activity was negatively correlated with depression severity (Avery et al., 2014). A role of the insula in MDD and treatment response seems evident, but it may come in different forms. For instance, four different functional regions of the insula were identified in a meta-analysis: activation in the anterior-ventral region was related to the social-emotional neural network; the mid-posterior region was activated in sensorimotor tasks; the central insula was activated in smell and taste; cognitive tasks activated the anterior-dorsal insula (Kurth et al., 2010).

Since, the present study utilizes the oddball task, the insula region mapping to a cognitive brain network is plausible to be involved. Activation of the insula makes sense in a paradigm in which salience, attention and working memory are relevant.

Indeed, the insula, together with the anterior cingulate, has been described as a hub in a salience network. Herein, it is involved in bottom-up detection of salient stimuli and facilitation of attention and working memory resources by switching between neural networks when a salient stimulus is perceived (Menon and Uddin, 2010).

Another explanation is that the effect that we found for the N1 amplitude is partly a reflection of the known relation between the LDAEP and treatment response (for instance (Gallinat et al., 2000; Mulert et al., 2007).

Why the prediction holds for only male responders to venlafaxine-XR needs further testing. Sex-specific and treatment-specific effects have been observed in earlier iSPOT-D reports though. For instance, it was found that, in women only, pre-treatment alpha-asymmetry predicted treatment response for escitalopram and sertraline, but not for venlafaxine-XR (Arns et al. 2015a). And in another study it was found that low frontal and low anterior cingulate cortex theta activity were associated with treatment response. Further post-hoc analysis revealed that the effect was mainly driven by one treatment, venlafaxine-XR (Arns et al. 2015b). So, in sum, sex certainly seems to have an effect in predicting treatment response by EEG measures in, and furthermore, the effect may depend on the type of medication. The size of the iSPOT-D sample allows for obtaining these nuances.

The study had several strengths and weaknesses. The multi-center nature of the study along with the large sample size, are strong aspects of the study. Furthermore, the use of two different analysis (peak-picking and eLORETA) methods that resulted in the same findings for P3 and N1 is a strength. A weakness is the already mentioned low spatial resolution of eLORETA (as implicated by the name: Low Resolution Tomography). It localizes brain activity exactly, but it does so on an approximate head model. Therefore, observed brain regions to which group differences in current source densities were located should be interpreted with caution. The analyses were also run for Remission, but this yielded no effects of N1 or P3 latencies/amplitudes (data not shown).

In conclusion, MDD patients compared to HC have low P3 amplitude with a small ES, unlikely to be of diagnostic utility. Furthermore, the N1 component from the ERP is a potential predictor for antidepressant treatment response to venlafaxine-XR, specifically in male MDD patients and with large effect sizes. Both main results also replicate results from previous small-scale studies, and the prediction results add emphasis on sex differences and localize differences to the insula. Regarding prediction of treatment outcome the observed results have limited clinical value. However, from an integrative perspective the N1 amplitude can provide more clinical value when it will be included in a sophisticated algorithm of multiple metric such as EEG and neuropsychology (Arns et al. (2015a, 2015b), Etkin et al., 2014). Future studies should investigate whether these markers, integrated with other markers, can improve prediction of treatment outcome. Furthermore, future treatment prediction studies in MDD should focus on sex as an important factor in analyses, and consider a priori stratification by sex.

Role of funding source

This study was funded by the Brain Resource Company Operations Ltd. The clinical trials registration identifier is NCT00693849. Data-analyses and writing of this manuscript were unconstrained and no financial support was involved in the data-analyses and writing of this manuscript.

Contributors

Author RvD managed the literature searches. Statistical analyses were carried out by RvD, MA and LW. Author RvD wrote the first draft of the manuscript. All authors contributed to and have approved the final manuscript.

Conflict of interest

RvD has no disclosures. MA reports research grants and options from Brain Resource Ltd. (Sydney, Australia), acted as a paid consultant for the United BioSource Corporation (UBC), Bracket, Mindmedia and Vivatech and is a co-inventor on 3 patent applications (A61B5/0402; US2007/0299323, A1; WO2010/139361 A1) related to EEG, neuromodulation and psychophysiology, but does not own these nor receives any proceeds related to these patents. KF reports research grants, options and shares from Brain Resource Ltd. PBF is supported by a NHMRC Practitioner Fellowship (606907). PBF has received equipment for research from MagVenture A/S, Medtronic Ltd, Cervel Neurotech and Brainsway Ltd and funding for research from Cervel Neurotech; CD has received support from Brain Resource, CNS response, St. Jude, Astra Zeneca, Takeda, Assurex and is a consultant for Pfizer and Genentech; EG is founder and receives income as Chief Executive Officer and Chairman for Brain Resource Ltd. He has stock options in Brain Resource Ltd.

Acknowledgments

We acknowledge the iSPOT-D Investigators Group, the contributions of iSPOT-D principal investigators at each site and the central management team (global coordinator Claire Day) and Chris Spooner and Donna Palmer for support in the data-analyses. This study was funded by the Brain Resource Company Operations Ltd. The clinical trials registration identifier is NCT00693849. Data-analyses and writing of this manuscript were unconstrained and no financial support was involved in the data-analyses and writing of this manuscript.

References

Arns, M., Bruder, G., Hegerl, U., Spooner, C., Palmer, D.M., Etkin, A., Fallahpour, K., Gatt, J.M., Hirshberg, L., Gordon, E., 2015a. EEG alpha asymmetry as a gender-specific predictor of outcome to acute treatment with different antidepressant medications in the randomized iSPOT-D study. Clin. Neurophysiol 10.1016/j.clinph.2015.05.032.

Arns, M., Drinkenburg, W.H.I.M., Fitzgerald, P.B., Kenemans, J.L., 2012. Neurophysiological predictors of non-response to rTMS in depression. Brain Stimul. 5, 569–576.

Arns, M., Etkin, A., Hegerl, U., Williams, L.W., DeBattista, C., Palmer, D.M., Fitzgerald, P.B., Harris, A., deBeuss, R., Gordon, E., 2015b. Frontal and rostral anterior cingulate (rACC) theta EEG in depression: implications for treatment outcome? Eur. Neuropsychopharmacol.

Arns, M., Gunkelman, J., Breteler, M., Spronk, D., 2008. EEG phenotypes predict treatment outcome to stimulants in children with ADHD. J. Integr. Neurosci. 7, 421–438.

Arns, M., Olbrich, S., 2014. Personalized medicine in ADHD and depression: use of pharmaco-EEG. Curr. Top Behav. Neurosci., 345–370 Springer-Verlag, Berlin.

Avery, J.A., Drevets, W.C., Moseman, S.E., Bodurka, J., Barcalow, J.C., Simmons, W.K., 2014. Major depressive disorder

is associated with abnormal interoceptive activity and functional connectivity in the insula. Biol. Psychiatry 76, 258–266.
Bruder, G.E., Kayser, J., Tenke, C.E., 2009. Event-related Brain Potentials in Depression: Clinical, Cognitive and Neurophysiologic Implications. Event-related Potential Components: The Ups and Downs of Brainwave Recordings. Oxford University Press, New York.
Bruder, G.E., Tenke, C.E., Stewart, J.W., Towey, J.P., Leite, P., Voglmaier, M., Quitkin, F.M., 1995. Brain event-related potentials to complex tones in depressed patients: relations to perceptual asymmetry and clinical features. Psychophysiology 32, 373–381.
Bruder, G.E., Tenke, C.E., Towey, J.P., Leite, P., Fong, R., Stewart, J.E., McGrath, P.J., Quitkin, F.M., 1998. Brain ERPs of depressed patients to complex tones in an oddball task: relation of reduced P3 asymmetry to physical anhedonia. Psychophysiology 35, 54–63.
Coffman, J.A., Torello, M.W., Lewis, C.E., Fortin, L.D., Martin, D. J., 1989. Event-related brain potentials in depressed patients treated with electroconvulsive therapy. Prog. Neuropsychopharmacol. Biol. Psychiatry 13, 687–692.
Davidson, R.J., Irwin, W., Anderle, M.J., Kalin, N.H., 2003. The neural substrates of affective processing in depressed patients treated with venlafaxine. Am. J. Psychiatry 160, 64–75.
Etkin, A., Patenaude, B., Song, Y.J., Usherwood, T., Rekshan, W., Schatzberg, A.F., Rush, A.J., Williams, L.M., 2014. A cognitive-emotional biomarker for predicting remission with antidepressant medications: a report from the iSPOT-D trial. Neuropsychopharmacology.
Fava, M., Rush, A., Trivedi, M.H., Nierenberg, A.A., Thase, M.E., Sackeim, H.A., Quitkin, F.M., Wisniewski, S., Lavori, P.W., Rosenbaum, J.F., Kupfer, D.J., 2003. Background and rationale for the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) study. Psychiatr. Clin. N. Am. 26, 457–494.
Friedman, D., 2003. Cognition and aging: a highly selective overview of event-related potential (ERP) data. J. Clin. Exp. Neuropsychol. 25, 702–720.
Gallinat, J., Bottlender, R., Juckel, G., Munke-Puchner, A., Stotz, G., Kuss, H.J., Mavrogiorgou, P., Hegerl, U., 2000. The loudness dependency of the auditory evoked N1/P2-component as a predictor of the acute SSRI response in depression. Psychopharmacology (Berl.) 148, 404–411.
Gangadhar, B.N., Ancy, J., Janakiramaiah, N., Umapathy, C., 1993. P300 amplitude in non-bipolar, melancholic depression. J. Affect. Disord. 28, 57–60.
Gratton, G., Coles, M.G., Donchin, E., 1983. A new method for off-line removal of ocular artifact. Electroencephalogr. Clin. Neurophysiol. 55, 468–484.
Işıntaş, M., Ak, M., Erdem, M., Oz, O., Ozgen, F., 2012. [Event-related potentials in major depressive disorder: the relationship between P300 and treatment response]. Turk. Psikiyatri. Derg. 23, 33–39.
Jaworska, N., De Somma, E., Blondeau, C., Tessier, P., Norris, S., Fusee, W., Smith, D., Blier, P., Knott, V., 2013. Auditory P3 in antidepressant pharmacotherapy treatment responders, non-responders and controls. Eur. Neuropsychopharmacol.
Kalayam, B., Alexopoulos, G.S., 1999. Prefrontal dysfunction and treatment response in geriatric depression. Arch. Gen. Psychiatry 56, 713–718.
Kawasaki, T., Tanaka, S., Wang, J., Hokama, H., Hiramatsu, K., 2004. Abnormalities of P300 cortical current density in unmedicated depressed patients revealed by LORETA analysis of event-related potentials. Psychiatry Clin. Neurosci. 58, 68–75.
Kemp, A.H., Hopkinson, P.J., Hermens, D.F., Rowe, D.L., Sumich, A. L., Clark, C.R., Drinkenburg, W., Abdi, N., Penrose, R., McFarlane, A., Boyce, P., Gordon, E., Williams, L.M., 2009. Fronto-temporal alterations within the first 200 ms during an attentional task distinguish major depression, non-clinical participants with depressed mood and healthy controls: a potential biomarker? Hum. Brain Mapp. 30, 602–614.
Kenemans, J.L., Kähkönen, S., 2011. How human electrophysiology informs psychopharmacology: from bottom-up driven processing to top-down control. Neuropsychopharmacology 36, 26–51.
Kurth, F., Zilles, K., Fox, P.T., Laird, A.R., Eickhoff, S.B., 2010. A link between the systems: functional differentiation and integration within the human insula revealed by meta-analysis. Brain Struct. Funct. 214, 519–534.
Liston, C., Chen, A.C., Zebley, B.D., Drysdale, A.T., Gordon, R., Leuchter, B., Voss, H.U., Casey, B.J., Etkin, A., Dubin, M.J., 2014. Default mode network mechanisms of transcranial magnetic stimulation in depression. Biol. Psychiatry 76, 517–526.
Marcus, M., Yasamy, M.T., Van Ommeren, M., Chisholm, D., Saxena, S., 2012. Depression: a global public health concern. World Health Organization Paper on Depression, 6–8.
Mayberg, H.S., Silva, J.A., Brannan, S.K., Tekell, J.L., Mahurin, R. K., McGinnis, S., Jerabek, P.A., 2002. The functional neuroanatomy of the placebo effect. Am. J. Psychiatry 159, 728.
McGrath, C.L., 2013. Toward a neuroimaging treatment selection biomarker for major depressive disorder. JAMA Psychiatry 70, 1.
Menon, V., Uddin, L.Q., 2010. Saliency, switching, attention and control: a network model of insula function. Brain Struct. Funct. 214, 655–667.
Mulert, C., Juckel, G., Brunnmeier, M., Karch, S., Leicht, G., Mergl, R., Möller, H.J., Hegerl, U., Pogarell, O., 2007. Prediction of treatment response in major depression: integration of concepts. J. Affect. Disord. 98, 215–225.
Nichols, T.E., Holmes, A.P., 2002. Nonparametric permutation tests for functional neuroimaging: a primer with examples. Hum. Brain Mapp. 15, 1–25.
Olbrich, S., Arns, S.M., 2013. EEG biomarkers in major depressive disorder: discriminative power and prediction of treatment response. Int. Rev. Psychiatry 25, 604–618.
Pascual-Marqui, R.D., 1999. Review of methods for solving the EEG inverse problem. International. J. Bioelectromagn. 1, 75–86.
Pascual-Marqui, R.D., 2002. Standardized low-resolution brain electromagnetic tomography (sLORETA): technical details. Methods Find Exp. Clin. Pharmacol. 24, 5–12.
Pascual-Marqui, R.D., 2007. Discrete, 3D Distributed, Linear Imaging Methods of Electric Neuronal Activity. Part 1: Exact, Zero Error Localization. arXiv preprint arXiv:0710.3341.
Pascual-Marqui, R.D., Esslen, M., Kochi, K., Lehmann, D., 2002. Functional imaging with low resolution brain electromagnetic tomography (LORETA): review, new comparisons, and new validation. Jpn. J. Clin. Neurophysiol. 30, 81–94.
Pascual-Marqui, R.D., Michel, C.M., Lehmann, D., 1994. Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. Int. J. Psychophysiol. 18, 49–65.
Paul, R.H., Gunstad, J., Cooper, N., Williams, L.M., Clark, C.R., Cohen, R.A., Lawrence, J.J., Gordon, E., 2007. Cross-cultural assessment of neuropsychological performance and electrical brain function measures: additional validation of an international brain database. Int. J. Neurosci. 117, 549–568.
Röschke, J., Wagner, P., 2003. A confirmatory study on the mechanisms behind reduced P300 waves in depression. Neuropsychopharmacology 28 (Suppl. 1), S9–S12.
Rush, A.J., Trivedi, M.H., Wisniewski, S.R., Nierenberg, A.A., Stewart, J.W., Warden, D., Niederehe, G., Thase, M.E., Lavori, P.W., Lebowitz, B.D., McGrath, P.J., Rosenbaum, J.F., Sackeim, H.A., Kupfer, D.J., Luther, J., Fava, M., 2006. Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D report. Am. J. Psychiatry 163, 1905–1917.
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, 22–33.
Spronk, D., Arns, M., Barnett, K.J., Cooper, N.J., Gordon, E., 2011. An investigation of EEG, genetic and cognitive markers of treatment response to antidepressant medication in patients with major depressive disorder: a pilot study. J. Affect. Disord. 128, 41–48.
Strik, W.K., Fallgatter, A.J., Brandeis, D., Pascual-Marqui, R.D., 1998. Three-dimensional tomography of event-related potentials during response inhibition: evidence for phasic frontal lobe activation. Electroencephalogr. Clin. Neurophysiol. 108, 406–413.
Urretavizcaya, M., Moreno, I., Benlloch, L., Cardoner, N., Serrallonga, J., Menchón, J.M., Vallejo, J., 2003. Auditory event-related potentials in 50 melancholic patients: increased N100, N200 and P300 latencies and diminished P300 amplitude. J. Affect. Disord. 74, 293–297.
van Dinteren, R., Arns, M., Jongsma, M., Kessels, R., 2014a. Combined frontal and parietal P300 amplitudes indicate compensated cognitive processing across the lifespan. Front. Aging Neurosci. 6, 294.
van Dinteren, R., Arns, M., Jongsma, M.L.A., Kessels, R.P.C., 2014b. P300 development across the lifespan: a systematic review and meta-analysis. PLoS One 9, e87347.
Vandoolaeghe, E., van Hunsel, F., Nuyten, D., Maes, M., 1998. Auditory event related potentials in major depression: prolonged P300 latency and increased P200 amplitude. J. Affect. Disord. 48, 105–113.
Ward, A.M., Schultz, A.P., Huijbers, W., Van Dijk, K.R., Hedden, T., Sperling, R.A., 2014. The parahippocampal gyrus links the default-mode cortical network with the medial temporal lobe memory system. Hum. Brain Mapp. 35, 1061–1073.
Williams, L.M., Rush, A.J., Koslow, S.H., Wisniewski, S.R., Cooper, N.J., Nemeroff, C.B., Schatzberg, A.F., Gordon, E., 2011. International Study to Predict Optimized Treatment for Depression (iSPOT-D), a randomized clinical trial: rationale and protocol. Trials 12, 4.
Williams, L.M., Simms, E., Clark, C.R., Paul, R.H., Rowe, D., Gordon, E., 2005. The test–retest reliability of a standardized neurocognitive and neurophysiological test battery: “neuromarker”. Int. J. Neurosci. 115, 1605–1630.