ORIGINAL INVESTIGATION

**Andrew F. Leuchter · Melinda Morgan · Ian A. Cook ·
Jennifer Dunkin · Michelle Abrams · Elise Witte**

Pretreatment neurophysiological and clinical characteristics of placebo responders in treatment trials for major depression

Abstract *Rationale:* High placebo response rates are a confound in treatment trials for major depressive disorder (MDD). A method for prospective identification of placebo responders could enhance the efficiency of clinical trials. *Objective:* The objective was to identify the neurophysiological, symptomatic, and cognitive characteristics of subjects who were likely to respond to placebo in clinical trials for MDD. *Methods:* Fifty-one subjects with MDD were treated in clinical trials with either fluoxetine ($n=24$) or venlafaxine ($n=27$) versus placebo. All subjects underwent pretreatment assessment with quantitative electroencephalographic (QEEG) power and cordance, as well as symptom ratings and neuropsychological testing. After a 1-week single-blind placebo lead-in, subjects were randomized to double-blind placebo controlled treatment with a medication or placebo. At the end of 8 weeks, the blind was broken and treatment response assessed. Response was defined by a final Hamilton Depression Rating Scale score of ≤10. *Results:* Of the medication-treated and placebo-treated subjects, 52% (13/25) and 38% (10/26) responded. Placebo responders had lower pretreatment frontocentral cordance in the theta frequency band than all other subjects ($P<0.006$) and medication responders in particular ($P<0.004$). Placebo responders also had faster cognitive processing time, as assessed by neuropsychological testing, and lower reporting of late insomnia ($P<0.03$). Exploratory examination of a multiple variable model for predicting placebo response was conducted using logistic regression, in which these three pretreatment measures accurately identified 97.6% of eventual placebo responders. *Conclusions:* These findings suggest that combined clinical, neurophysiological, and cognitive assessments of prospective subjects for clinical trials may be useful for identifying MDD subjects who are likely to show robust response to placebo. Prospective validation of these results in a larger, independent sample of subjects is necessary to establish the reliability and usefulness of this method for prospective identification of placebo responders.

Keywords Depression · Medication treatment · Placebo response · Quantitative EEG · Cordance · Cognitive processing · Insomnia

From the Laboratory of Behavioral Pharmacology, UCLA Neuropsychiatric Institute, and the Department of Psychiatry and Biobehavioral Sciences, David Geffen School of Medicine at UCLA, Los Angeles, CA.

A. F. Leuchter (✉)
UCLA NPI,
760 Westwood Plaza, Suite 37-452,
Los Angeles, CA, 90024-1759, USA
e-mail: afl@ucla.edu
Tel.: +1-310-8250207
Fax: +1-310-8257642

M. Morgan
UCLA NPI,
760 Westwood Plaza, Suite 37-432,
Los Angeles, CA, 90024-1759, USA

I. A. Cook
UCLA NPI,
760 Westwood Plaza, Suite 37-451,
Los Angeles, CA, 90024-1759, USA

J. Dunkin
UCLA NPI,
760 Westwood Plaza, Suite 37-439,
Los Angeles, CA, 90024-1759, USA

M. Abrams
UCLA NPI,
760 Westwood Plaza, Suite 37-431,
Los Angeles, CA, 90024-1759, USA

E. Witte
26314 NE 37th Avenue,
Ridgefield, WA, 98642, USA

Introduction

Between 30% and 50% of depressed subjects in clinical trials have a substantial reduction in symptoms during treatment with placebo (Beecher 1955; Shapiro and Shapiro 1997; Walsh et al. 2002; Fava et al. 2003). While the placebo response is both an interesting and useful clinical phenomenon, it bedevils antidepressant

drug discovery. "Failed trials" (those in which drug fails to separate from placebo) are common and cost the pharmaceutical industry hundreds of millions of dollars annually (Enserlink 1999; Robinson and Rickels 2000). Reduction in the rates of placebo response could enhance the efficiency of screening potential antidepressant compounds, decrease the needed size (and therefore the cost) of clinical trials, and potentially decrease the number of failed trials (Fava et al. 2003).

Control over placebo response rates has proved to be elusive. One strategy has been to alter the design of the clinical trial, but there is little or no agreement among studies regarding which factors are crucial to control (Fava et al. 2003; Michelson et al. 1999; Zimbroff and Mendez 2002). Another strategy has been the single-blind placebo lead-in, in which subjects receive placebo for a brief period with the expectation that eventual placebo responders (PRs) will show a response during the lead-in (Landin et al. 2000). Only a small minority of subjects actually responds to placebo during a 1-week lead-in, however, and there is no overall reduction in placebo response rates (Reimherr et al. 1989; Trivedi and Rush 1994; Faries et al. 2001).

Other lines of work have aimed at identifying the characteristics of individuals who constitute likely placebo responder (LPR) subjects. The most successful work in this regard is that of Quitkin and colleagues (Quitkin et al. 1993; Quitkin 1999; Stewart et al. 1998) who examined the time-course of improvement during a clinical trial. While this "pattern analysis" method is powerful, it can identify PRs only at the end of a clinical trial.

A more useful method would be prospective identification of LPR subjects so that they could be screened out of clinical trials using a selective enrollment strategy. Demographic factors such as subject age, gender, education level, or occupation have not proved to be reproducible predictors of response (Shapiro and Shapiro 1997). Depressive symptoms may have usefulness in predicting which subjects will respond to placebo (Peselow et al. 1992). McGrath et al. (2000) found that those subjects with more severe neurovegetative symptoms on the Hamilton Depression Rating Scale (Ham-D) were less likely to respond to placebo.

We previously reported changes in brain function during treatment in depressed subjects who responded to placebo (Leuchter et al. 2002). In the present study, we examined this same group of individuals to determine whether there were pretreatment factors that would prospectively identify subjects who were eventual PRs. We hypothesized that data from three different domains would be useful for identifying LPR subjects prior to treatment. These domains were neurophysiological (quantitative electroencephalographic; QEEG), neuropsychological (results from a cognitive testing battery), and symptomatic (neurovegetative items from the Ham-D). The neurophysiological data included QEEG power and cordance, a measure that has moderately strong associations with cerebral metabolism and perfusion (Leuchter et al. 1994a,b, 1999). In contrast to our previous work that focused on a limited number of brain regions (Leuchter et al. 2002), in the present study we examined brain function from all recording electrodes to determine whether any brain region would show pretreatment differences in brain function. Neuropsychological data were collected for seven spheres of cognitive function. Four individual symptom items from the 17-item Ham-D were selected based on previous findings regarding neurovegetative symptoms (early, middle, and late insomnia, as well as appetite) (McGrath et al. 2000). Items from each domain were examined in an exploratory multiple variable model to determine whether a combination of data from these three domains might be useful for prospective identification of LPR subjects.

Materials and methods

Subjects

Subjects were recruited from community advertisement and from the outpatient clinics of the UCLA Neuropsychiatric Hospital, and were enrolled in one of two 9-week, double-blind, placebo-controlled treatment studies. These studies were conducted sequentially and independently over a 24-month period: the first study utilized fluoxetine 20 mg (n=24) and the second venlafaxine 150 mg (n=27) as total daily medication doses. The UCLA IRB approved all experimental procedures, and all procedures were performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki. Written

Table 1 Demographic characteristics of subjects. There were no significant clinical or demographic differences between the subjects enrolled in the fluoxetine or venlafaxine trials

	Fluoxetine study n=24	Venlafaxine study n=27	Placebo responders n=10	Placebo non-responders n=16	Medication responders n=13	Medication non-responders n=12
Age (years±SD)	40.3±11.5	42.6±12.5	38.2±15.2	43.2±9.6	41.2±11.7	42.3±13.2
Proportion of female subjects	16/24	16/27	4/10	10/16	10/13	8/12
Proportion of subjects with positive family history of depression	12/24	20/27	5/10	12/16	7/13	7/12
Baseline Ham-D	21.8±4.2	22.4±3.1	20.7±3.3	22.6±3.6	21.5±3.0	23.3±4.3
Final Ham-D	12.3±8.3	13.1±6.2	6.1±3.4	17.9±4.5	6.0±3.1	18.4±4.6
Total number of depressive episodes	2.3±1.1	2.0±1.3	2.2±1.4	1.9±0.7	1.9±0.9	2.6±1.7

Table 2 Tests comprising the seven spheres of cognitive function

No.	Cognitive sphere	Variables
1	Information processing speed	Trails A (Reitan 1958)
		Stroop Color Word Task A and B (Stroop 1935)
		Digit Symbol Test (Wechsler 1981)
2	Executive functioning	Trails B (Reitan 1958)
		Wisconsin Card Sort Test (Lezak 1983)
		Auditory Consonant Trigrams (Stuss et al. 1985)
		Controlled Oral Word Association Test (Benton 1968)
		Stroop Color Word Trail (Stroop 1935)
3	Language	Boston Naming Test (Kaplan et al. 1983)
4	Verbal memory	Rey Auditory Verbal Learning Test (Rey 1964)
5	Non-verbal memory	WMS-R Visual Reproduction (Wechsler 1987)
		Rey-Osterreith Complex Figure (Rey 1941)
6	Basic attention	Digit Span (Woodcock and Johnson 1977)
7	Visuoperceptual ability	WAIS-R Block Design (Wechsler 1981)
		Rey-Osterrieth Complex Figure Copy (Rey 1941)
		Benton Facial Recognition (Benton et al. 1978)

informed consent was obtained after experimental procedures were explained fully to the subjects, and prior to their inclusion in the study.

All subjects were over 21 years of age, met DSM-IV criteria for a major depressive episode, and had 17-item Hamilton Depression Rating Scale (Ham-D) scores ≥16 (Hamilton 1960). Diagnoses were established using the Structured Clinical Interview for DSM-IV. Subjects were excluded if they previously had failed treatment with the antidepressant being studied, had a history of suicidal ideation, or suffered from any medical illness or received any medication known to significantly affect brain function. Demographic and clinical characteristics of the subjects are shown in Table 1.

Assessment instruments

The severity of depression was assessed at baseline and throughout the study period using the 17-item Ham-D (Hamilton 1960). The total score of the Ham-D was used as the entry criterion for the study, as well as to determine treatment response. In order to examine differences in neurovegetative symptoms across response groups, items for early, middle, and late insomnia, as well as appetite were selected for further analysis.

Cognitive function was assessed using a battery of tests that broadly measured cognitive performance in seven spheres: (1) information processing speed, (2) executive functioning, (3) language, (4) verbal memory, (5) non-verbal memory, (6) basic attention, and (7) visuoperceptual ability. The specific tests in each sphere are detailed in Table 2. Cognitive testing data were available only on 42 of the 51 subjects; the remaining 9 subjects were unable to complete testing.

Experimental procedures

Fifty-three subjects received 1 week of single-blinded placebo lead-in treatment. Two subjects met response criteria (Ham-D ≤10) at the end of this week and were removed from the studies. The remaining 51 subjects then were randomized to receive 8 weeks of double-blind treatment with either placebo or the active medication. Subjects enrolled in the fluoxetine trial received 20 mg/day for the entire 8 weeks. Those enrolled in the venlafaxine trial began at 37.5 mg/day and increased the dose by 37.5 mg every 3 days, until a dose of 75 mg b.i.d. was attained after 1 week. They continued at a dose of 150 mg/day for the remaining 7 weeks. To preserve blinding, the number of placebo pills was increased on the same schedule as the medication.
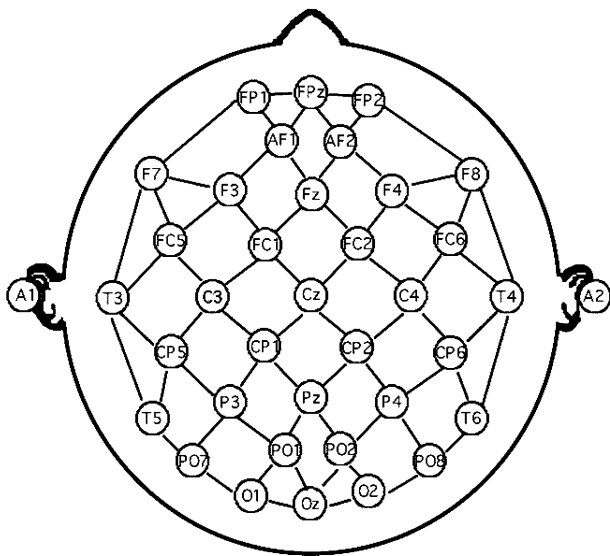
After randomization, research staff evaluated subjects at 2 days and at weekly intervals thereafter. In addition to symptom evaluation using the Ham-D, subjects had brief supportive psychotherapy at each visit, consisting of 15–25 min of unstructured counseling and assistance in problem solving by a research nurse. This support was required by the IRB to address concerns about dispensing placebo as the sole treatment for patients with major depression. Response was defined as Ham-D ≤10 after 8 weeks of double-blind treatment. At this time, blind was broken and subjects were classified as medication responders (MRs), PRs, medication non-responders (MNRs), or placebo non-responders (PNRs).

QEEG techniques

QEEG data were examined from a recording performed at the time of enrollment, before any treatments were administered. Recordings were performed using the QND system (Neurodata, Inc.; Pasadena, CA) while subjects rested in the eyes-closed, maximally alert state in a sound-attenuated room with subdued lighting, using procedures previously described in detail (Leuchter et al.

1999; Cook et al. 1998, 1999). Electrodes were placed with an electrode cap (ElectroCap; Eaton, OH) using an extended International 10–20 System with 35 recording electrodes (Fig. 1). Eye movements were monitored with right infraorbital and left outer canthus electrodes. Data were collected using a Pz reference montage and were digitized at 256 samples/channel/s, with a high-frequency filter of 50 and a low-frequency filter of 0.3 Hz. Data were reformatted by amplitude subtraction to construct a linked-ears reference montage, and then were reviewed by a technician who was blinded to subject identity, treatment condition, and clinical status. The technician examined the record carefully for eye movement, muscle, or other artifacts, and selected the first 20–32 s of artifact-free data for processing. This amount of data may be used to obtain reliable frequency spectra (Leuchter et al. 1992, 1999; Brenner et al. 1995). A fast Fourier transform was used to calculate absolute power (the intensity of energy in a frequency band in microvolts squared) in each of four frequency bands (0.5–4, 4–8, 8–12, and 12–20 Hz).

Cordance values next were calculated for each electrode site in each of the four frequency bands. Cordance is a measure derived from QEEG power and has a moderately strong association with cerebral perfusion (as assessed by simultaneous O^{15} positron emission tomography); this association is superior to that seen for conventional QEEG power measurements in each frequency band (Leuchter et al. 1999). Cordance is calculated using a three-step algorithm that normalizes power across both electrode sites and frequency bands. This algorithm has been defined in detail elsewhere (Leuchter et al. 1999) and may be summarized briefly as follows. First, absolute power values are reattributed to each individual electrode by averaging power from all bipolar electrode pairs

sharing that electrode (Fig. 1). This electrode referencing method is similar to the Hjorth transformation, except that the present method averages power from neighboring electrode pairs, whereas the Hjorth transformation averages voltage amplitudes. We previously reported that electrode referencing based on power averaging provides a stronger association between surface-measured EEG and perfusion of underlying brain than either the linked-ears reference or the conventional Hjorth transformation (Cook et al. 1998). Based on reattributed absolute power, relative power (the percentage of the total energy from all bands concentrated in a single band) is also calculated. Second, absolute and relative power values undergo spatial normalization within each frequency band using a z-score transformation, yielding z-scores for each electrode site s and frequency band f ($A_{\text{norm}(s,f)}$ and $R_{\text{norm}(s,f)}$, respectively). Third, the z transformed power scores are summed to yield cordance values.

For each individual over the course of treatment, we calculated power and cordance values for individual electrodes (Fig. 1). In contrast with our previous study, in which we focused on certain specific brain regions (Leuchter et al. 2002), we examined data from each recording electrode. We limited our analysis to the theta frequency band (4–8 Hz), because previous work from this and other laboratories has indicated that energy in the theta band is associated most strongly with treatment outcomes in depression (Cook et al. 1999; Ulrich et al. 1984, 1994).

Data analysis

Demographic and clinical differences among the four response groups (MR, MNR, PR, and PNR) were examined first. Relationships between categorical independent variables and a single continuous dependent variable were examined using analysis of variance (ANOVA). Chi-square analyses were used to explore possible relationships when both the independent and dependent variables were categorical.

A dichotomous treatment response variable was created to look at PRs versus all other study participants (MR, MNR, and PNR). In order to conduct a more focused examination of the responders, we also created a dichotomous response variable including only the PRs and MRs.

We hypothesized that each of the three domains of data at baseline (neurophysiological, neuropsychological, and neurovegetative symptoms) would be useful for identifying PRs. Multivariate analysis of variance (MANOVA) was performed first to determine whether there were differences among the groups of subjects in multiple dependent variables (QEEG power and cordance) in the neurophysiological domain. For the neuropsychological testing domain, a series of MANOVAs was performed to determine whether there was a difference among the groups in any cognitive sphere (if a sphere consisted of a single measure, ANOVA was used instead of MANOVA). MANOVA was utilized because it corrects for multiple comparisons by maintaining the groupwise error rate at

Fig. 1 Electrode montage. The 35 scalp electrodes from the extended International 10–20 System. "Neighboring" electrodes are linked by line segments, to denote bipolar channels that were used for averaging in the reattributional montage. All electrode pairs sharing a common electrode were averaged to obtain a reattributed power for that electrode (e.g., for electrode C3, power from the pairs FC5-C3, FC1-C3, CP5-C3, and CP1-C3 were averaged)

0.05. Box's M was used to test the assumption of equality of covariance matrices and Levene's test was used to assess equality of error variance across groups. In domains where the omnibus MANOVA was significant, univariate *F*-tests were then used to further examine the effect of the specific dependent variables that contributed to the overall effect. Pillai's Trace was used because it is the most robust test statistic when unequal sample sizes are examined.

After the initial examination of the data, exploratory logistic regression was used to develop a multivariable model to estimate the probability of placebo response. Neurophysiological and neuropsychological variables that showed between-group differences in the MANOVAs as well as the four neurovegetative items from the Ham-D were used as the independent variables. Forward stepwise regression was used to select variables with the strongest predictive value to enter the final model based on a classification cutoff of 0.5. Nagelkerke's *R* squared was used to assess the strength of association between the independent and dependent variables for the final logistic regression model. The chi-square model was used to assess the improvement in fit when the independent variables were in the model versus the null model. Logistic regression coefficients were assessed using the Wald statistic to test the significance of the individual variable, while holding constant all other independent variables in the model.

Results

Clinical outcomes

A comparison of the subjects from the two medication trials showed that their demographic and clinical characteristics, medication and placebo response rates, mean final Ham-D scores, and dropout rates were not significantly different (Table 1). Data from the two trials, therefore, were pooled for further analyses.

Overall, 52% of the subjects (13/25) receiving antidepressant medication responded to treatment, while 38% of those receiving placebo (10/26) responded; the difference in response rates was not statistically significant. Both responder groups had lower final Ham-D scores than nonresponders, but no group could be distinguished from any other based on pretreatment levels of depression, number of prior episodes, family history of depression, or any demographic characteristics (Table 1). Both responder groups had similar rates of decline in depression and ended with substantially lower depression scores than nonresponders (Fig. 2).

Neurophysiological data

MANOVA for QEEG absolute and relative power showed no difference among the groups of subjects for any of the 35 electrodes. For QEEG cordance, the omnibus *F* was 2.58 with *P*=0.023. The follow-up univariate tests indicated the only significant group differences were in
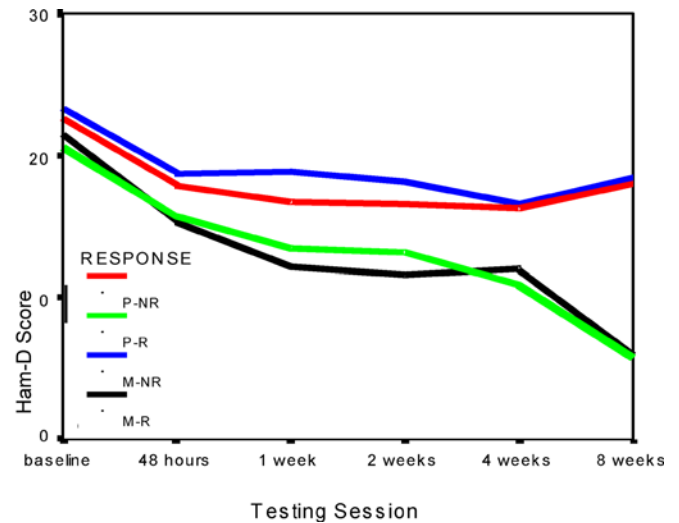
Fig. 2 Changes in mean Ham-D scores over time for the four outcome groups. There were no differences at any time point between the mean scores of the medication responders (MR) and placebo responders (PR), although both groups of responders were significantly different from both non-responder groups at all time points (*P*<0.05) except baseline

electrodes Af1 and Af2. Because these two electrodes were immediately adjacent to each other over the frontocentral region, mean values for these two electrodes were averaged to create a frontocentral cordance value for each treatment response group. Frontocentral cordance was significantly lower in PR subjects than in all other subjects in general (*P*=0.006, two tailed, *t*=2.87) and in the MR subjects in particular (*P*=0.004, two tailed, *t*=3.29) (Table 3).

Table 3 Group mean differences in frontocentral cordance. Cordance values from Af1 and Af2 electrodes were averaged to obtain measure of frontocentral cordance. *PR* placebo responder, *MR* medication responder

Frontocentral cordance	*n*	Mean	SD	*t*	*P*
PR	10	−1.21	1.14	2.87	0.006
All other subjects	41	−0.001	1.20		
PR	10	−1.21	1.14	3.29	0.004
MR	13	0.26	0.99		

Table 4 Analyses of variance for neuropsychological tests for each sphere of cognitive function. Only information processing speed showed a difference among the groups of subjects

Cognitive sphere	*F* statistic (Pillai's Trace)	*P*
Information processing speed	2.85	0.037
Executive functioning	1.33	0.260
Language[a]	0.161	0.690
Verbal memory	0.876	0.507
Non-verbal memory	0.327	0.918
Basic attention[a]	0.138	0.712
Visuoperceptual ability	0.796	0.504

[a] ANOVA was used in these analyses because the sphere of function consisted of only one variable

Neuropsychological data

Results of the MANOVA and ANOVA tests for baseline differences among the groups of subjects are shown in Table 4. Only the sphere of information processing speed was significant, within which further univariate tests indicated that the difference was attributable to the PR subjects' performance on Digit Symbol Test, which was faster than that for all other groups combined (Table 5). PRs also performed better, though not significantly better, than all other groups combined on Trails A and the Stroop Color Word tests. There were no significant differences between the subset of PR and MR subjects alone (data not presented).

Neurovegetative symptom data

There were no pretreatment differences among the treatment outcome groups in overall severity of depression as measured by the total Ham-D scores. Univariate testing for the four neurovegetative items, however, revealed that the PR subjects reported significantly lower levels of late insomnia than all other subjects combined (t=2.22, two tailed, P=0.03). There was no significant difference between PR and MR subjects alone, and no difference among the groups in the other three Ham-D items.

Logistic regression

Frontocentral cordance, Digit Symbol Test, late, middle, and early insomnia, as well as appetite were offered as candidate variables for logistic regression. Using the forward stepwise method of entry, the final model included frontocentral cordance, the Digit Symbol Test, and late insomnia. Nagelkerke's R squared value was 0.653, demonstrating a moderately strong relationship between the independent variables and the dependent variable. The significance level of the omnibus test of the model coefficients was 0.001, indicating a strong improvement in fit when the independent variables were in the model versus the model with only the constant. The P values for the individual coefficients based on the Wald statistic were all significant (P<0.03) (Table 6). This model correctly classified 97.6% of the subjects into their treatment response category (Table 7). A separate logistic regression carried out for PR versus MR subjects alone revealed that frontocentral cordance was the only variable

Table 5 Group mean difference on the Digit Symbol Test. Neuropsychological testing data were available on 42 of 51 subjects and showed a significant difference between placebo responders (PRs) and all other subjects combined. There was no difference between PR or medication responder (MR) subjects alone

Cognitive test item	N	Mean	SD	T	P
Digit Symbol					
PR	8	12.75	1.91	−2.26	0.03
All other subjects	34	11.09	1.86		

selected for the model, with 75% of the subjects classified correctly.

Discussion

These results suggest that it may be possible, on the basis of pretreatment measures of brain function, symptom severity, and cognitive performance, to identify prospectively those subjects who are likely to exhibit a placebo response at the end of a clinical trial. These subjects showed substantially lower cordance in the frontocentral region, had slightly less late insomnia, and had slightly faster cognitive processing times than did those subjects who did not show a placebo response. The differences in brain function were significant whether the PR subjects were compared with all other response groups combined, or only with those subjects who responded to medication. The fact that PR subjects differed specifically and significantly from MR subjects suggests that it may be possible to identify prospectively a selective group of LPR subjects for future studies.

The finding that frontocentral brain activity is associated with responsiveness to placebo treatment is consistent with prior literature implicating the anterior cingulate region in mediating the response to antidepressant medication (Mayberg et al. 1997). The frontocentral recording electrodes from which we detect brain functional characteristics of PRs overlie the anterior cingulate region, and activity from these electrodes in the theta-freqency band has been reported to reflect anterior cingulate metabolism (Asada et al. 1999; Ishii et al. 1999).

Previous studies have suggested that PRs have less severe illness than PNRs. The PR subjects in this study did not differ significantly in terms of overall severity of depression from other subjects, but did have slightly less severe late insomnia, a finding consistent with the literature (McGrath et al. 2000; Thase et al. 1993). The finding that overall severity of depression did not differ between the responder groups indicates that baseline brain functional differences between PR and other subjects cannot be explained solely on the basis of illness severity.

To our knowledge, cognitive processing speed never has been examined relative to the placebo response in depression. Previous studies have reported that MRs are more likely to show normal cognition than MNRs, who frequently show deficits on tests of executive function (Dunkin et al. 2000; Alexopoulos et al. 2000), although PR subjects did not differ from any other responder group in executive or other measures of cognitive function. The present results suggest that those subjects who respond to placebo may have the least amount of cognitive slowing of all subjects with major depression.

It is important to note several limitations of the current study. First, the number of PRs on whom we report is relatively small. These results, therefore, should be interpreted with caution, and await replication on a larger number of subjects. Second, although the proportion of subjects responding to medication was greater than the number responding to placebo, this difference was not

Table 6 Logistic regression. Variable(s) entered on step 1, frontocentral cordance; variable (s) entered on step 2, late insomnia; variable(s) entered on step 3, Digit Symbol

		B	SE	Wald	df	Sig.	Exp(B)
Step 1	Frontocentral cordance	−0.860	0.369	5.422	1	0.02	0.42
	Constant	−2.050	0.579	12.527	1	0.00	0.13
Step 2	Frontocentral cordance	−1.100	0.449	5.994	1	0.014	0.33
	Late insomnia	−1.377	0.696	3.920	1	0.05	0.25
	Constant	−0.842	0.720	1.369	1	0.24	0.43
Step 3	Frontocentral cordance	−1.429	0.646	4.887	1	0.03	0.24
	Late insomnia	−2.885	1.216	5.626	1	0.02	0.06
	Digit Symbol	1.278	0.574	4.950	1	0.03	3.59
	Constant	−14.857	6.530	5.177	1	0.02	0.00

Table 7 Classification table. Classification matrix based on 42 of 51 subjects who had complete data (including neuropsychological testing). *PR* placebo responder

	Predicted		Percentage accuracy of prediction
Observed	All other subjects	PR	
All others	34	0	100.0
PR	1	7	87.5
Overall percentage			97.6

statistically significant. The lack of a significant difference in response rates is an increasingly frequent occurrence in clinical trials (Walsh et al. 2002) and may reflect a lower than usual medication response rate (52%), a higher than usual placebo response rate (38%), or a combination of these factors. One factor contributing to the significant placebo response rate in this study may have been the supportive psychotherapy offered to all subjects. Although this interaction with the subjects was not substantially different from the support offered to subjects in most antidepressant clinical trials, it may have helped engender the placebo response in these subjects (Fava et al. 2003) and could limit the generalizability of the findings of this study. Third, it is likely that some of the subjects in the MR group actually constitute PRs (i.e., received medication but improved for reasons unrelated to medication treatment). We have no method to identify the PRs who may be embedded within the MR group. PRs had significantly lower prefrontal cordance than the group of MRs, however, so their removal could actually increase the between-group difference. Fourth, the rate of correct classification with logistic regression must be interpreted with caution both because of the limited sample size and because of the absence of an independent sample of subjects on whom to test the classification variables. The results of the logistic regression should be interpreted primarily as demonstrating the potential value of the multimodal approach to classification of PRs, and await replication on an independent sample of subjects.

It would be ideal if restrictive entry criteria could be employed to exclude LPR subjects from phase-IIb clinical trials in order to reduce placebo response rates. Reduced placebo response rates could reduce the number of subjects required for these early trials and improve the efficiency with which trials could detect medication

effectiveness (Fava et al. 2003). If the findings of this study can be replicated, it might be possible to utilize a combination of neurophysiological, neuropsychological, and symptomatic measures as a screening method to identify and exclude LPR subjects from these medication development studies. It is important to note that, while this type of subject screening and restrictive entry criteria could be useful for rapid identification of lead compounds for antidepressant drug development, it would not at this juncture be applicable to phase-III trials. Restrictive entry criteria in trials of medication efficacy could result in selection of study populations that would not be representative of patients in general practice.

Whether the multimodal approach described here would be practical for phase-IIb trials would depend ultimately on a cost–benefit analysis: would the burden of the assessments proposed here constitute a barrier to subject enrollment, and would the expense of this testing battery be offset by the cost savings of smaller trials? In our experience, the assessment described here could be completed in less than 2 h and poses a modest additional cost and minimal burden to subjects. A more complete cost–benefit analysis, however, would be an important component of future studies. The preliminary results reported here are encouraging, however, and suggest that a combination of brain functional, cognitive, and symptomatic measures may be a useful strategy for reducing placebo response rates in medication development trials. This approach, if replicated by additional studies, could be a useful component of early clinical trials.

References

Alexopoulos GS, Meyers BS, Young RC, Kalayam B, Kakuma T, Gabrielle M, Sirey JA, Hull J (2000) Executive dysfunction and long-term outcomes of geriatric depression. Arch Gen Psychiatry 57:285–290

22

Asada H, Fukuda Y, Tsunoda S, Yamaguchi M, Tonoike M (1999) Frontal midline theta rhythms reflect alternative activation of prefrontal cortex and anterior cingulate cortex in humans. Neurosci Lett 274:29–32

Beecher HK (1955) The powerful placebo. JAMA 159:1602–1606

Benton AL (1968) Differential behavioral effects in frontal lobe disease. Neuropsychologia 6:53–60

Benton AL, Varney NR, Hamsher KD (1978) Visuospatial judgment. A clinical test. Arch Neurol 35:364–367

Brenner RP, Ulrich RF, Reynolds CF (1995) EEG spectral findings in healthy elderly men and women—sex differences. EEG Clin Neurophys 94:1–5

Cook IA, O'Hara R, Uijtdehaage S, Mandelkern M, Leuchter AF (1998) Assessing the accuracy of topographic EEG mapping for determining local brain function. EEG Clin Neurophys 107:404–414

Cook IA, Leuchter AF, Witte E, Abrams M, Uijtdehaage SHJ, Stubbeman W, Rosenberg-Thompson S, Anderson-Hanley C (1999) Neurophysiologic predictors of treatment response to fluoxetine in major depression. Psychiatry Res 85:263–273

Dunkin JJ, Leuchter AF, Cook IA, Kasl-Godley JE, Abrams M, Rosenberg-Thompson S (2000) Executive dysfunction predicts nonresponse to fluoxetine in major depression. J Affect Disord 60:13–23

Enserlink M (1999) Psychopharmacology: can the placebo be the cure? Science 284:238–240

Faries DE, Heiligenstein JH, Tollefson GD, Potter WZ (2001) The double-blind variable placebo lead-in period: results from two antidepressant clinical trials. J Clin Psychopharm 21:561–568

Fava M, Evins AE, Dorer DJ, Schoenfeld DA (2003) The problem of the placebo response in clinical trials for psychiatric disorders: culprits, possible remedies, and a novel study design approach. Psychother Psychosom 72:115–127

Hamilton M (1960) A rating scale for depression. J Neurol Neurosurg Psychiatry 23:56–62

Ishii R, Shinosaki K, Ukai S, Inouye T, Ishihara T, Yoshimine T, Hirabuke N, Asada H, Kihara T, Robinson SE, Takeda M (1999) Medial prefrontal cortex generates frontal midline theta rhythm. Neuroreport 10:675–679

Kaplan E, Goodglass H, Weintraub S (1983) Boston Naming Test. Lea & Febiger, Philadelphia, PA

Landin R, DeBrota DJ, DeVries TA, Potter WZ, Demitrack MA (2000) The impact of restrictive entry criterion during the placebo lead-in period. Biometrics 56:271–278

Leuchter AF, Newton TF, Cook IA, Walter DO, Rosenberg-Thompson S, Lachenbruch PA (1992) Changes in brain functional connectivity in Alzheimer's-type and multi-infarct dementia. Brain 115:1543–1561

Leuchter AF, Cook IA, Lufkin RB, Dunkin J, Newton TF, Cummings JL, Mackey K, Walter D (1994a) Cordance: a new method for assessment of cerebral perfusion and metabolism using quantitative electroencephalography. Neuroimage 1:208–219

Leuchter AF, Cook IA, Mena I, Dunkin J, Cummings JL, Newton T, Migneco O, Lufkin R, Walter D, Lachenbruch P (1994b) Assessment of cerebral perfusion using quantitative EEG cordance. Psychiatry Res Neuroimaging 55:141–152

Leuchter AF, Uijtdehaage SH, Cook IA, O'Hara R, Mandelkern M (1999) Relationship between brain electrical activity and cortical perfusion in normal subjects. Psychiatry Res Neuroimaging 90:125–140

Leuchter AF, Cook IA, Witte EA, Morgan M, Abrams M (2002) Changes in brain function of depressed subjects during treatment with placebo. Am J Psychiatry 159:122–129

Lezak MD (1983) Neuropsychological Assessment. Oxford University, New York

Mayberg HS, Brannan SK, Mahurin RK, Jerabek PA, Brickman JS et al (1997) Cingulate function in depression: a potential predictor of treatment response. Neuroreport 8:1057–1061

McGrath PJ, Stewart JW, Petkova E, Quitkin FM, Amsterdam JD, Fawcett J, Reimherr FW, Rosenbaum JF, Beasley CM Jr (2000) Predictors of relapse during fluoxetine continuation or maintenance treatment of major depression. J Clin Psychiatry 61:518–524

Michelson D, Holland P, Tepner R, Pemberton C (1999) Double-blind vs. single-blind placebo lead-in periods during panic disorder efficacy trials. Poster presented at the 39th Annual NCDEU Meeting. Boca Raton, FL

Peselow ED, Sanfilipo MP, Difiglia C, Fieve RR (1992) Melancholic/endogenous depression and response to somatic treatment and placebo. Am J Psychiatry 159:1848–1854

Quitkin FM (1999) Placebos, drug effects and study design: a clinician's guide. Am J Psychiatry 156:829–836

Quitkin FM, Stewart JW, McGrath PJ (1993) Columbia atypical depression. A subgroup of depressives with better response to MAOI than to tricyclic antidepressants or placebo. Br J Psychiatry 163:30–34

Reimherr FW, Ward MF, Byerley WF (1989) The introductory placebo washout: a retrospective evaluation. Psychiatry Res 30:191–199

Reitan RM (1958) Validity of the trail making test as an indicator of organic brain damage. Percept Mot Skills 8:271–276

Rey A (1941) L'examen psychologique dans les cas d'encephalopathie traumatique. Arch Psychol 28:286–340

Rey A (1964) L'examen clinique en psychologie. Pressess Universitiaires de France, Paris

Robinson DS, Rickels K (2000) Concern about clinical drug trials. J Clin Psychopharmacol 20:593–596

Shapiro AK, Shapiro E (1997) The Powerful Placebo: from Ancient Priest to Modern Physician. Johns Hopkins University, Baltimore

Stewart JW, Quitkin FM, McGrath PJ, Amsterdam J, Fava MD, Fawcett J, Reimherr F, Rosenbaum J, Beasley C, Roback P (1998) Use of pattern analysis to predict differential relapse of remitted patients with major depression during one year of treatment with fluoxetine or placebo. Arch Gen Psychiatry 55:334–343

Stroop JR (1935) Studies of interference in serial verbal reactions. J Exp Psychol 12:643–662

Stuss DT, Ely P, Hugenholtz H, Richard MT, LaRochelle S, Poirier CA, Bell I (1985) Subtle neuropsychological deficits in patients with good recovery after closed head injury. Neurosurgery 17:41–47

Thase ME, Simons AD, Reynolds CF (1993) Psychobiological correlates of poor response to cognitive behavior therapy: potential indications for antidepressant pharmacotherapy. Psychopharmacol Bull 29:293–301

Trivedi MH, Rush H (1994) Does a placebo run-in or a placebo treatment cell affect the efficacy of antidepressant medications? Neuropsychopharmacology 11:33–43

Ulrich G, Renfordt E, Zeller G, Frick K (1984) Interrelation between change in the EEG and psychopathology under pharmacotherapy for endogenous depression: a contribution to the predictor question. Pharmacopsychiatry 17:178–183

Ulrich G, Haug HJ, Fahndrich E (1994) Acute versus chronic EEG effects in maprotiline- and in clomipramine-treated depressive inpatients and the prediction of therapeutic outcome. J Affect Disord 32:213–217

Walsh BT, Seidman SN, Sysko R, Gould M (2002) Placebo response in studies of major depression: variable, substantial and growing. JAMA 287:1840–1847

Wechsler D (1981) Manual for the Wechsler Adult Intelligence Scale—revised. Psychological Corporation, New York

Wechsler D (1987) Manual for the Wechsler Memory Scale—revised. Psychological Corporation, New York

Woodcock RW, Johnson MB (1977) Woodcock-Johnson Psycho-Educational Battery. DLM Teaching Resources, Allen, Texas

Zimbroff DL, Mendez, G (2002) Patient and rater education about expectations in clinical trials (PREECT): an approach to reducing placebo response rates and improving signal detection in psychopharmacology trials. Poster presented at the 42nd Annual NCDEU Meeting. Boca Raton, FL