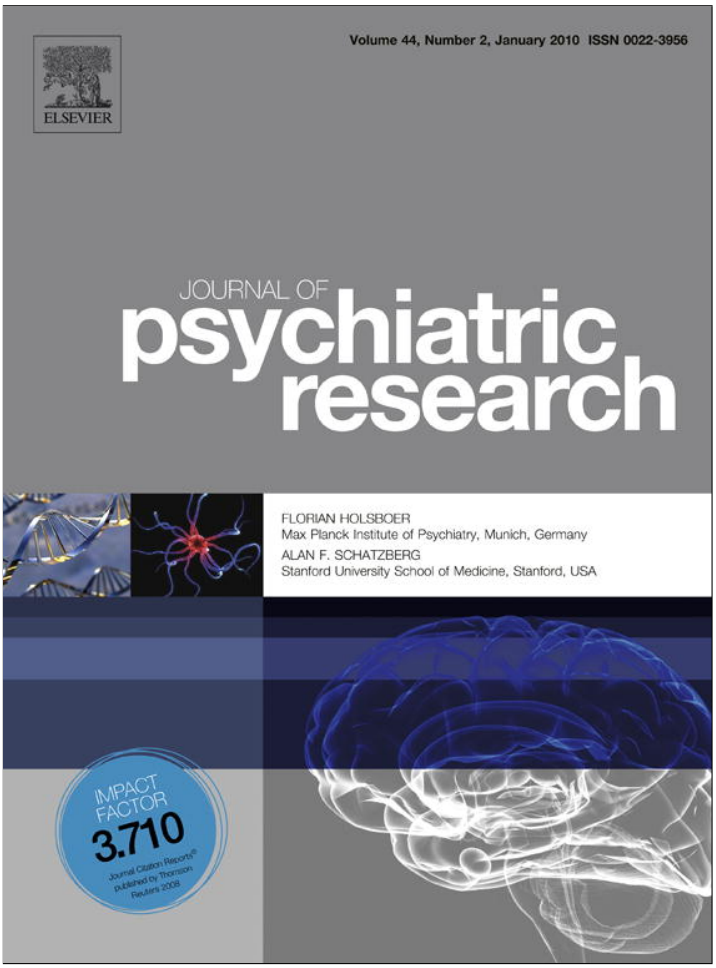

Journal of Psychiatric Research 44 (2010) 90–98

Contents lists available at ScienceDirect

Journal of Psychiatric Research

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

Antidepressant response trajectories and quantitative electroencephalography (QEEG) biomarkers in major depressive disorder

Aimee M. Hunter [a,*], Bengt O. Muthén [b], Ian A. Cook [a], Andrew F. Leuchter [a]

[a] Semel Institute for Neuroscience and Human Behavior at UCLA, Department of Psychiatry and Biobehavioral Sciences, David Geffen School of Medicine at UCLA, Westwood Plaza, Rm. 37-359, Los Angeles, CA 90024-1759, United States
[b] UCLA Graduate School of Education and Information Studies, Los Angeles, United States

ARTICLE INFO

Article history:
Received 24 March 2009
Received in revised form 11 June 2009
Accepted 17 June 2009

Keywords:
Depression
EEG
Cordance
Antidepressants
Statistical models
Growth mixture modeling

ABSTRACT

Individuals with Major Depressive Disorder (MDD) vary regarding the rate, magnitude and stability of symptom changes during antidepressant treatment. Growth mixture modeling (GMM) can be used to identify patterns of change in symptom severity over time. Quantitative electroencephalographic (QEEG) cordance within the first week of treatment has been associated with endpoint clinical outcomes but has not been examined in relation to patterns of symptom change. Ninety-four adults with MDD were randomized to eight weeks of double-blinded treatment with fluoxetine 20 mg or venlafaxine 150 mg ($n = 49$) or placebo ($n = 45$). An exploratory random effect GMM was applied to Hamilton Depression Rating Scale (Ham-D_{17}) scores over 11 timepoints. Linear mixed models examined 48-h, and 1-week changes in QEEG midline-and-right-frontal (MRF) cordance for subjects in the GMM trajectory classes. Among medication subjects an estimated 62% of subjects were classified as responders, 21% as non-responders, and 17% as symptomatically volatile—i.e., showing a course of alternating improvement and worsening. MRF cordance showed a significant class-by-time interaction ($F_{(2,41)} = 6.82$, $p = .003$); as hypothesized, the responders showed a significantly greater 1-week decrease in cordance as compared to non-responders (mean difference = −.76, Std. Error = .34, $df = 73$, $p = .03$) but not volatile subjects. Subjects with a volatile course of symptom change may merit special clinical consideration and, from a research perspective, may confound the interpretation of typical binary endpoint outcomes. Statistical methods such as GMM are needed to identify clinically relevant symptom response trajectories.

1. Introduction

Treatment with antidepressant medications is associated with significant improvements in clinical symptoms of Major Depressive Disorder (MDD), as well as improvements in functional status and quality of life. However there is marked heterogeneity in clinical outcomes and there are no reliable means of predicting such varied outcomes for the individual patient.

Clinical research in MDD often utilizes a single primary endpoint measure to assess response/non-response. Although this approach is useful, information is lost regarding variability in patterns of response. Individuals differ, not only with respect to magnitude of response, but also time-to-response and stability of response. For example, although clinically significant symptom reduction often is not observed until 4–6 weeks after starting an antidepressant medication (Donovan et al., 1994; Nierenberg et al., 2000; Quitkin et al., 1996), marked improvement may occur as early as the first two weeks (Papakostas et al., 2007; Posternak and Zimmerman, 2005; Stassen et al., 2007) or as late as eight or more weeks after initiating treatment (Trivedi et al., 2006). Regarding symptom stability, whereas most patients show monotonic improvement, some experience transient clinical worsening and/or treatment-emergent adverse events in the first few months of treatment (Cusin et al., 2007; Perahia et al., 2008). This is an important subgroup because of their poorer long-term prognosis (Cusin et al., 2007; Perahia et al., 2008) increased rates of discontinuation (Beasley et al., 2000; Chelben et al., 2001; Kaplan, 1997), and the discomfort and potential danger of exacerbation of symptoms.

Statistical growth modeling offers a formalized approach to identifying symptom response patterns. Repeated measurement linear mixed (multilevel) modeling incorporates outcome measures at all timepoints to estimate a 'trajectory shape' response pattern over time. An advantage of examining whole trajectories is that the outcome does not rely on a single timepoint that includes possible day-to-day fluctuation, and instead draws on measures from several weeks, considering trends. However, the linear mixed model with its random effects is not sufficient to capture fundamentally different trajectory shapes. In this regard, growth mixture modeling (GMM), a multilevel modeling technique that

* Corresponding author. Tel.: +1 310 206 2237; fax: +1 310 825 7642.
E-mail address: amhunter@ucla.edu (A.M. Hunter).

0022-3956/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jpsychires.2009.06.006

incorporates features of cluster analysis, is particularly advantageous for analyses of clinical trials because it allows for different treatment effects in different trajectory classes (Muthén and Shedden, 1999; Muthén and Muthén, 2000; Muthén et al., 2002; Muthén and Asparouhov, 2008; Muthen et al., 2008; Muthén and Brown, in press). GMM can be applied to longitudinal data to identify latent "classes" or patterns of change in symptom severity over time. This technique is newly emerging in the literature on treatment outcomes and, although there are several reports utilizing GMM to examine symptom changes during psychotherapeutic interventions for MDD (Cuijpers et al., 2005; Stulz et al., 2007; Stulz and Lutz, 2007), none yet has specifically focused on pharmacotherapy outcomes.

Changes in brain function early in the course of antidepressant treatment have been related to simple endpoint clinical outcomes. Quantitative electroencephalography (QEEG) imaging measures within the first week of treatment have been shown to predict end-of-trial outcomes with over 70% accuracy (Hunter et al., 2007) but have never before been examined in relation to patterns of symptom change. Early changes in QEEG cordance, in particular, have been associated with endpoint response and remission in double blind placebo-controlled trials (Cook et al., 2002; Leuchter et al., 2005; Cook et al, in press) and have been found to predict antidepressant response across independent research institutions (Bares et al., 2007, 2008). Cordance incorporates both absolute and relative power and has been shown to have a stronger association with cerebral perfusion as measured by Positron Emission Tomography (PET) than either absolute or relative power alone (Leuchter et al., 1999).

The motivation to examine symptom change patterns and EEG characteristics in the same cohort of subjects is that the generation of response classes based upon symptom trajectories might be validated by evidence of neurophysiologic differences (Muthén, 2004). As such, the objectives of the present study were to: (1) identify patterns of change in depressive symptoms over the course of eight-week antidepressant trials for MDD, and (2) to examine early changes in QEEG cordance for subjects identified in each symptom trajectory class. We hypothesized that early regional changes in QEEG cordance that previously have distinguished endpoint responders vs. non-responders might also distinguish GMM-identified response trajectories.

2. Method

2.1. Subjects and design

Data were collected from 94 adults with MDD who participated in one of three placebo-controlled antidepressant trials that included QEEG cordance imaging. Recruitment mechanisms as well as inclusion and exclusion criteria were identical for the three protocols. Subjects were recruited through outpatient clinics and community advertisement and met MDD diagnostic criteria using a structured interview for DSM-IV (First et al., 1995), with a 17-item Hamilton Depression Rating Scale (Ham-D_{17}; (Hamilton, 1960)) score $\geqslant$16. Subjects were excluded if they previously had failed to benefit from treatment with the antidepressant being studied, if they had a history of suicide attempt, or if they suffered from any medical illness or received any medication known to significantly affect brain function. Enrolled subjects included 58 females and 36 males with a mean age of 41.7 ± 13.3 years, a mean baseline Ham-D_{17} score of 22.0 ± 3.7 points, and a mean final Ham-D_{17} score of 12.1 ± 7.6 points. Table 1 shows clinical and demographic characteristics of the sample for the three trials. The pooled trials had a common design except for the active medication used. A one-week placebo lead-in was followed by eight weeks of double-blind randomized treatment with active drug (fluoxetine 20 mg. in Study 1; venlafaxine 150 mg. in Studies 2 and 3) or placebo (Table 1). Concomitant use of psychotropic medications (e.g., sedatives or hypnotics) was prohibited during the trial and all subjects had been free of psychotropic medications for at least two weeks prior to beginning the study. Ham-D_{17} scores were obtained at every visit: baseline, end of placebo lead-in, 48 h after start of randomized treatment, and weekly throughout eight weeks of randomized treatment. QEEG measures were examined at pretreatment baseline, and at 48-h and 1-week timepoints after treatment assignment. The UCLA Institutional Review Board reviewed and approved the protocols, and, after complete description of the study to the subjects, written informed consent was obtained.

2.2. QEEG method and cordance calculations

EEG recordings were performed using 35 recording electrodes positioned with an electrode cap (ElectroCap, Inc., Eaton, OH) using an extended International 10-20 System (Fig. 1). Recordings were obtained while subjects rested in the eyes-closed, maximally alert state in a sound-attenuated room with subdued lighting, using the QND system (Neurodata, Inc., Pasadena, CA) with a Pz reference montage. Eye movements were monitored using right infraorbital and left outer canthus electrodes. Data were digitized online at 256 samples per channel per second with a high-frequency filter of 70 Hz and a low-frequency filter of 0.3 Hz, and were reformatted by amplitude subtraction to construct a bipolar electrode pair montage. An EEG technologist blinded to subject identity and treatment condition selected for processing the first 20–32 s of artifact-free data. An independent blinded technologist confirmed

Table 1
Clinical and demographic characteristics of subjects in randomized placebo-controlled trials using fluoxetine or venlafaxine as the active medication; no significant differences were observed across the three trials.

	Study 1 *N*=28	Study 2 *N*=33	Study 3 *N*=33	Test	Two-tailed *p*-value
Active medication	Fluoxetine: *n* = 14	Venlafaxine: *n* = 17	Venlafaxine: *n* = 18		
Placebo	*n* = 14	*n* = 16	*n* = 15		
Age (years)	42.4(12.6)	44.7(14.0)	38.1(12.5)	$F_{(2,91)} = 2.13$	.13
Gender ratio					
Female:male	19:9	21:12	18:15	Chi-Sq. = 1.33, df = 2	.54
Treatment history					
None prior:prior history	16:11	14:19	11:18	Chi-Sq. = 2.84, df = 2	.24
Initial Ham-D_{17}	22.1(4.2)	22.3(3.1)	21.5(3.8)	$F_{(2,91)} = .51$	.60
Final Ham-D_{17}	12.5(8.5)	13.0(6.4)	10.8(8.1)	$F_{(2,68)} = .58$	.56
Ham-D_{17} Response (50% improvement)					
Responder: Non-responder					
Medication	6:6	5:6	9:4	Chi-Sq. = 1.59, df = 2	.45
Placebo	6:4	5:9	5:6	Chi-Sq. = 1.39, df = 2	.50

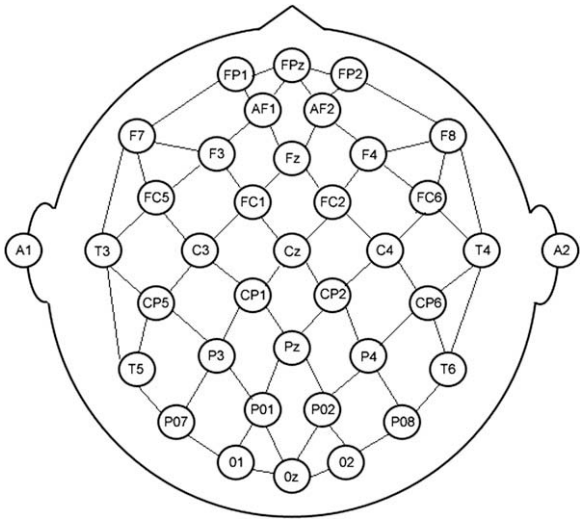

Fig. 1. Extended International 10–20 montage used for recording. Lines between electrodes indicate nearest-neighbor bipolar electrode pairs that were averaged for reattributed power calculations (Cook et al., 2002).

the selection prior to processing. 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 Hz, 4–8 Hz, 8–12 Hz, and 12–20 Hz).

Cordance values were calculated from conventional QEEG absolute and relative power measures in each of the four frequency bands for each electrode site. This three-step procedure is described elsewhere in greater detail (Leuchter et al., 1999) and has been employed in a number of prior reports (e.g., Cook et al., 2002; Leuchter et al., 2008). First, EEG power values were computed using a re-attributional electrode montage because this montage affords a higher correlation between EEG measures and PET measures of cerebral perfusion than other montages (Cook et al., 1998). Second, the absolute and relative power values were *z*-transformed to measure deviation from the mean values for each electrode site *s* in each frequency band *f* for that recording, yielding $A_{\mathrm{norm}(s,f)}$ and $R_{\mathrm{norm}(s,f)}$, respectively. Third, these *z*-scores were summed to yield a cordance "intensity" value, *Z*, for each electrode in each frequency band where $Z_{(s,f)} = A_{\mathrm{norm}(s,f)} + R_{\mathrm{norm}(s,f)}$. Analyses for this report focused on changes in theta-band (4–8 Hz) cordance in as measured from electrodes overlying the midline-and-right-frontal (MRF) region (electrodes Af2, F4, F8, Fp2, Fpz, Fz). This regional marker was previously associated with brain functional effects (Leuchter et al., 2008) and clinical effects (Cook et al., in press; Leuchter et al., 2005) of antidepressant medication.

2.3. Data analysis

2.3.1. GMM

GMM analyses were conducted using Mplus version 5 (Muthén and Muthén, 1998–2008). We applied a piece-wise GMM (Muthén and Muthén, 2000) focusing on Ham-D_{17} scores over 11 timepoints in separate analyses for subjects randomized to medication versus placebo. The first piece of the model included Ham-D_{17} measures at baseline and at end of placebo lead-in (i.e., visits prior to start of medication). To capture potential volatility in a flexible way for the early stages of the trial, the second piece incorporated the 48-h measurement and weeks 1–2, with a third piece corresponding to weeks 3–8. For the first piece, a linear model was used with a random intercept only, while for the second and third piece a quadratic model was used with random intercepts and slopes. Given that subjects are randomly equivalent during the first piece, all 94 subjects were used to estimate the parameters of this piece. For timepoints after start of randomized treatment, data from medication subjects only ($n = 49$) were included. A second analysis was performed for placebo subjects ($n = 45$), examining the same 3-piece model. The analysis did not specify *a priori* trajectory shapes but was exploratory in nature, allowing shapes to be found using a flexible growth model. The GMM analyses were carried out using maximum-likelihood estimation (MLE) with the expectation-maximization (EM) algorithm.

In both analyses, both a 2- and a 3-class GMM were used. This is based on the notion that both a responder and non-responder trajectory class needs to be represented, as well as a possible further class with a volatile development. For the analysis involving the medication subjects the 2-class GMM found a responder class and a volatile class of 19%. The 3-class solution showed a responder, a non-responder, and a volatile class. The volatile class of the 3-class GMM had very similar trajectory shape and prevalence as the 2-class GMM. The 3-class GMM was therefore chosen for further investigation. It should be noted that the customary statistical approach of using the Bayesian information criterion (BIC) to aid the decision on the number of classes does not work well for a sample of this size (Nylund et al., 2007). For the 3-class analysis involving the medication subjects, the GMM resulted in a good classification quality as reflected by entropy of 0.80 (Ramaswamy et al., 1993). For the 3-class placebo group analysis the classification quality was even better with an entropy of 0.89.

2.3.2. Analyses of clinical/demographic variables and brain function (i.e. QEEG cordance) among GMM outcome classes

Analyses examining clinical/demographic characteristics and brain functional measures among the GMM classes were conducted using SPSS version 16. Baseline clinical/demographic characteristics and brain function (cordance) were assessed among the GMM outcome classes using Chi-square for categorical variables (gender) and ANOVA for continuous variables (illness severity, MRF cordance). Regarding early changes in brain function, we focused on changes in MRF theta-band cordance within the first week of antidepressant treatment. Prior work has shown that decreases in MRF cordance at one week reflect neurophysiologic effects of antidepressant medication (Leuchter et al., 2008) and predict end-of-trial clinical response (Cook et al., in press; Leuchter et al., 2005). We examined MRF cordance change using a linear mixed model (random intercept model) with full MLE with time as the within-subjects factor (change at 48 h or 1 week) and GMM class as the between-subjects factor. We were interested in assessing the potential main effect of trajectory class on cordance or a potential class-by-time interaction as either of these effects would suggest differences in the course of early neurophysiologic change among subjects in the various clinical outcome classes. In addition, we specifically hypothesized that medication subjects classified as responders would show a greater decrease in MRF cordance at one week as compared to subjects in a non-responder class.

3. Results

3.1. End-of-trial outcomes

Of 94 subjects across the three trials, 71 completed through the primary endpoint (week 8). Using a standard response criterion of 50% improvement on the Ham-D_{17}, 16 of 35 (46%) of placebo subjects responded and 20 of 36 (56%) medication subjects responded; there was no significant difference in response rates between medication vs. placebo subjects. Table 1 shows 50% improvement response rates for medication and placebo across the three trials.

3.2. GMM symptom trajectory shape classes

3.2.1. Medication subjects

Fig. 2a shows estimated Ham-D17 symptom response trajectories for three classes of medication response. The three outcomes groups can be described as: 'responders' (62%), 'non-responders' (21%), and 'symptomatically volatile' (17%). The 'responder' group shows a symptom trajectory that would meet typical end-of-trial response criteria of 50% improvement or final Ham-D17 score $\leqslant$10. The non-responder class shows fairly steady improvement through week 3 before leveling off at a Ham-D17 score close to 16. This pattern shows partial response but with a high degree of residual symptom severity. Lastly, the symptomatically volatile class trajectory shows a fluctuating course of alternating improvement and worsening that can be viewed in four segments: sharp improvement 48 hours after start of drug, a 4-point worsening between 48 hours and week 2, improvement from weeks 2 through week 6 (to achieve the same symptom level as responders), and finally a 5-point worsening between weeks 5 and 8. Fig. 2b shows observed individual trajectories of those subjects who were classified in the symptom volatility class. Fig. 2c summarizes mean (±standard error of the mean) Ham-D total scores at each timepoint for subjects classified in the three GMM clinical outcome groups.

3.2.2. Placebo subjects

Estimated symptom response trajectories for placebo subjects are shown in Fig. 3. The trajectory shape for Class 1 (34.6%) is distinguished from Class 2 (43.3%) and Class 3 (22.1%) by a higher estimated Ham-D_{17} score at baseline. This group shows the steepest initial improvement through 48 h after beginning randomized treatment; however the trajectory is fairly flat from 48 h through week 8 with scores remaining within a 3 1/2 point spread. Class 2 shows a substantial 6-point improvement through the first week of randomized treatment before reaching a plateau Ham-D_{17} score of about 14. Class 3, shows the slowest initial rate of improvement through week 3 but then shows a rapid curvilinear decline in symptom severity reaching an estimated Ham-D_{17} score of 8 at week 8. In comparison to the Class 2 response trajectory in medication subjects, none of the placebo response trajectory shapes appears to show a similar degree of symptom volatility.

3.2.3. Baseline Clinical/Demographic and QEEG Characteristics of Medication Subjects among GMM Classes

Medication subjects among the three GMM outcome classes did not differ significantly with respect to age ($F_{(2,46)} = .37$, $p = .69$) or gender (Chi square = .39, df = 2, $p = 83$) but did differ with respect to baseline illness severity ($F_{(2,46)} = 18.90$, $p < .0001$). Subjects in the non-responder group had the highest mean baseline Ham-D_{17} score (27.25, SD = 3.06) and this was significantly higher than the mean score for responders (20.92, SD = 2.47; $t_{(31)} = -5.97$, $p < .0001$) and subjects in the symptom volatility group (20.40, SD = 2.97, $t_{(11)} = -3.97$, $p < .002$).

Fig. 4a illustrates baseline cordance values for subjects classified according to the three outcome groups. Because outcome groups did not show baseline differences in the MRF region ($F_{(2,43)} = .80$, $p = .46$), change-from-baseline MRF cordance values were calculated for each subject at the 48-h and 1-week timepoints.

3.3. Regional QEEG cordance changes among the GMM classes

Fig. 4b and c show whole head topographic changes in QEEG cordance for medication subjects at 48 h and at 1 week. Fig. 5 illustrates early changes in MRF cordance for the three symptom trajectory classes. Subjects in the responder class showed a decrease in MRF cordance beginning at 48 h and continuing through week 1. MRF brain functional changes in the non-responder group and the symptom volatility group appeared to track closely with each other; these outcome classes were characterized by large decreases in cordance at 48 h and a return toward baseline levels at week 1. The linear mixed model examining MRF cordance changes at 48 h and at 1 week among the three GMM medication outcome classes did not find a main effect of group ($F_{(2,45)} = .36$, $p = .70$) or time ($F_{(1,41)} = 3.58$, $p = .07$) but did find a significant group-by-time interaction ($F_{(2,41)} = 6.82$, $p = .003$). Based on our model we performed two *a priori* hypothesis tests to assess whether a greater decrease in MRF cordance at 1 week would be associated with response, first as compared to non-response, and second as compared to volatile response. Tests using a t statistic revealed a significant difference in MRF cordance change at week 1 between the responder group and the non-responder group (mean difference = −.76, Std. Error = .34, $df = 73$, $p = .03$), but not between the responder group and the symptom volatility group (mean difference = −.33, Std. Error = .35, $df = 63$, $p = .36$). As a basis for comparison we also examined the same linear mixed model applied to subjects randomized to placebo; consistent with our hypotheses, the model examining MRF cordance changes among the *placebo* response trajectories did not find a significant of effect group or time or a significant interaction.

3.4. Post Hoc analyses

3.4.1. MRF cordance Co-varying for baseline illness severity

Because we found a significant difference in baseline illness severity among the three GMM medication outcome classes, we examined a linear mixed model that included change in MRF cordance and baseline Ham-D_{17} score as a covariate. As in the model examining only MRF cordance, the model yielded a significant group-by-time interaction ($F_{(2,41)} = 6.81$, $p = .003$). However, the difference in MRF cordance change at week 1 between the responder and non-responder class did not reach significance (mean difference = −.61, Std. Error = .43, df = 61, $p = .16$) when controlling for baseline severity indicating that there is some shared variability between these predictors.

3.4.2. Overlap between subjects identified as belonging to the symptom volatility class using GMM, and subjects classified as responders using endpoint outcome

GMM identified 7 of 49 (14%) medication subjects as belonging to the symptom volatility class. Of these volatile subjects, six completed the medication trial through week 8. Using a standard endpoint response criterion of $\geqslant$50% improvement on the Ham-D_{17}, two of the six volatile completers would be considered responders.

4. Discussion

This study found three distinct patterns of change in symptom severity over eight weeks of antidepressant treatment for which there was some evidence of neurophysiologic differences. GMM analyses identified 'responder,' 'non-responder,' and 'symptom volatility' trajectory shapes estimated to comprise 62%, 21%, and 17% of subjects, respectively. The responder and non-responder patterns are consistent with conceptualizations of monotonic improvement over the course of antidepressant treatment with greater effectiveness for some patients (responders) and lesser effectiveness for other (non-responders). In contrast, the symptom volatility trajectory suggests that a small subgroup of patients exhibits a highly fluctuating course symptom severity with alternating periods of improvement and worsening. Changes in the QEEG marker, MRF cordance, were different across 48-h and 1-

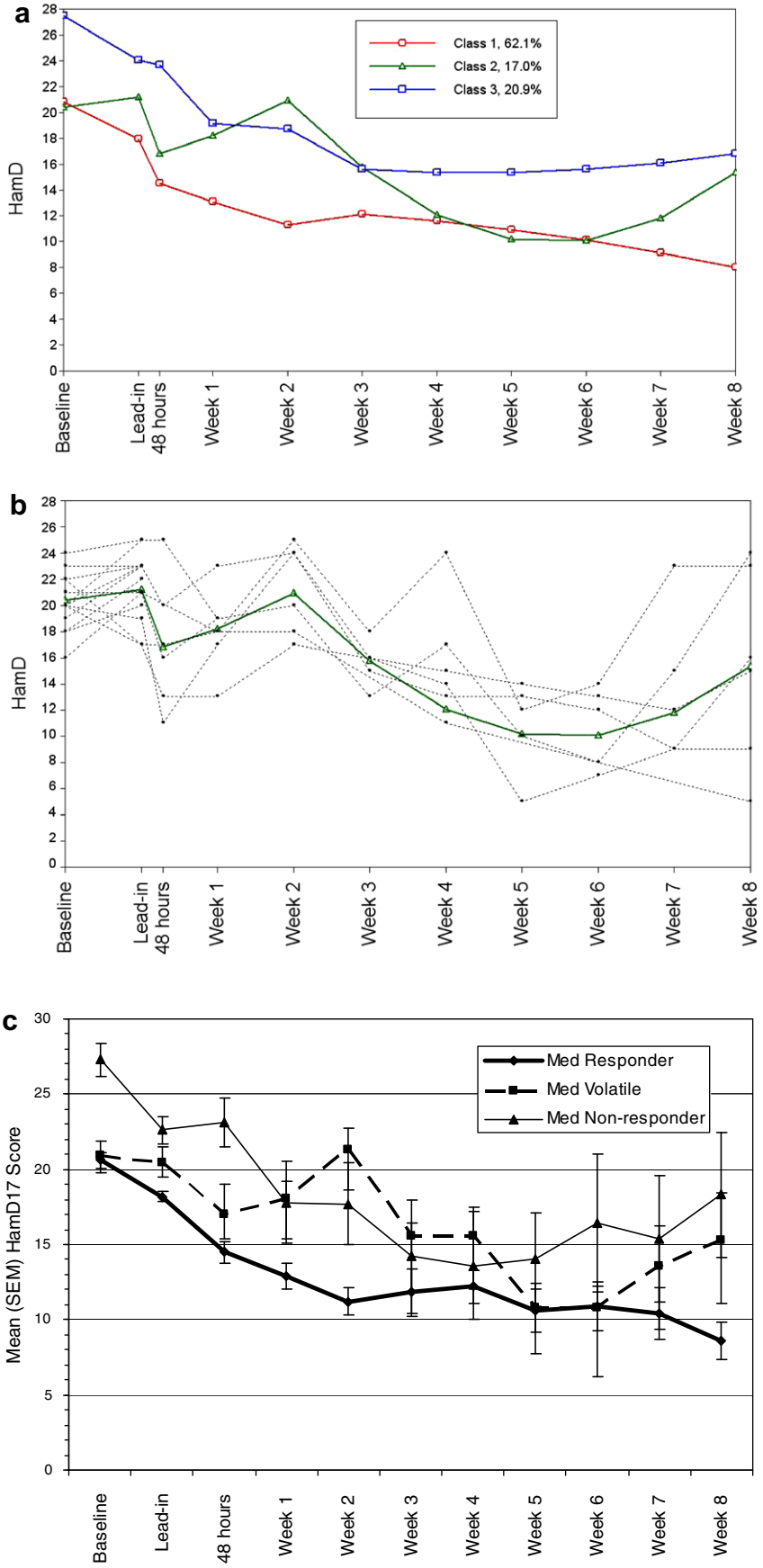

Fig. 2. (a) Estimated Ham-D_{17} means for three classes of MDD subjects randomized to antidepressant medication (top); (b) Individual trajectories for those subjects classified as Class 2 (middle); (c) Mean (±standard error of the mean) HamD total scores for medication subjects assigned to each GMM class.

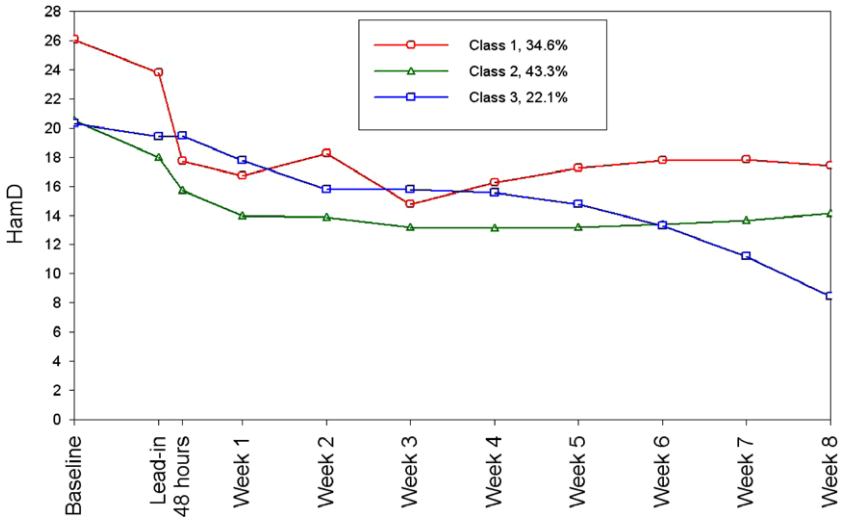

Fig. 3. Estimated Ham-D_{17} means for three groups of MDD subjects randomized to placebo.

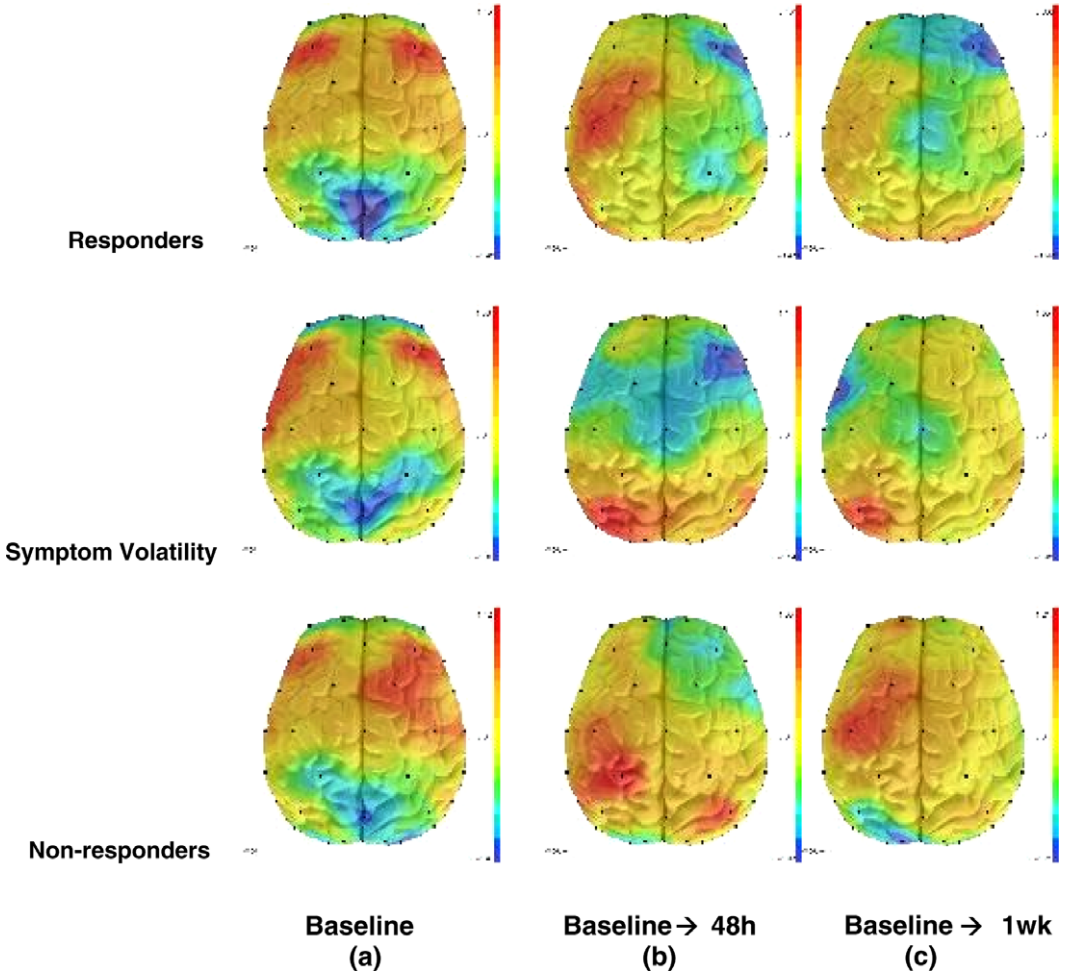

Fig. 4. Brain maps showing baseline (column a) and change-from-baseline theta cordance (columns b and c) for medication subjects classified as responders (62%), symptomatically volatile (17%), or non-responders (21%).

week timepoints as a function of symptom trajectory class among subjects randomized to antidepressant medication.

Interestingly, although non-responders were shown to have significantly greater illness severity at baseline, there was no difference in baseline Ham-D_{17} scores between responder and symptom volatility groups. Thus, at the outset of treatment, subjects who would later show stable steady improvement versus an erratic symptom course were indistinguishable in terms of overall symptom severity as measured the Ham-D_{17} total score.

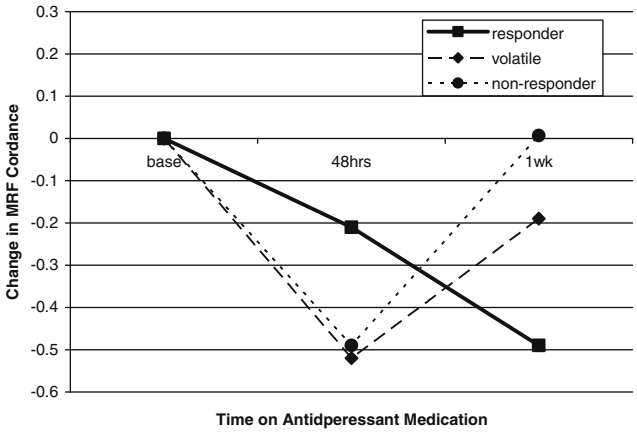

Fig. 5. MRF cordance changes from baseline for the three GMM outcome classes.

The symptom volatility group may be of special clinical interest. First, this subgroup may represent patients whose depression is especially difficult to treat and/or monitor. Prior reports have suggested that periods of clinical worsening in the early stages of antidepressant treatment may portend poorer long-term outcome (Cusin et al., 2007; Perahia et al., 2008). Moreover, subjects who exhibit early worsening are at greater risk for discontinuation of treatment (Beasley et al., 2000; Chelben et al., 2001; Kaplan, 1997). Second, the cause of symptom volatility in this subgroup is unknown. One possibility is that the fluctuating symptoms represent an unstable placebo-like response. This idea would be in keeping with conceptions put forth by Quitkin and colleagues that early improvement followed by worsening (i.e., an inverted U pattern of response) reflects placebo effects rather than "true drug" effects (Quitkin et al., 1991, 1987, 1984). However, the pattern found here is polyphasic, and, in the present study, placebo subjects did not exhibit a similarly volatile pattern. Third, it is possible that this symptom pattern is detecting heterogeneity in depressive illness. It is curious that the rapid fluctuation of symptom changes in this group is similar to the reported reaction of patients with bipolar disorder (BPD) to antidepressant treatment, namely fluctuation of symptoms and induction of "cycling." Although we have no direct evidence, one could speculate that some subjects diagnosed with MDD actually were suffering from a bipolar-spectrum illness (Akiskal, 1993), such as BPD type II, which is characterized primarily by depressive episodes. For these individuals, antidepressant treatment could possibly induce rapidly oscillating symptoms (American Psychiatric Association, 2002; Sachs et al., 2000).

The observation that there may be a subgroup with unstable response also has implications for research on markers of antidepressant 'response/non-response' or 'remission/non-remission' when those classifications are derived using endpoint outcomes. Whether outcome is predicted by QEEG or other imaging markers, genetic markers, or clinical and demographic characteristics, the ability to accurately predict a dichotomous outcome partly depends upon the stability of that outcome measure. A 'symptom volatility' pattern is clearly an unfavorable clinical outcome yet use of a single endpoint criterion leaves the dichotomous responder/non-responder classification of symptomatically unstable subjects to chance. Thus any finding based upon unstable symptom change outcomes introduces a source of variance into the outcome of clinical trials that may be difficult to detect. If an estimated 17% of subjects were to fall into this category, it could introduce a sizeable margin of error into any clinical trial. Unless subjects with unstable symptom changes are identified as non-responders, their inclusion in endpoint analyses may obscure ability to test both the effectiveness of treatment and the predictive capability of a biomarker for response.

The MRF cordance biomarker, previously associated with antidepressant effects and endpoint response or remission, was found to differ between medication subjects classified according to responder versus non-responder outcome trajectories. Consistent with prior observations showing an association between decreases in MRF theta cordance within the first week of antidepressant treatment, and end-of-trial clinical improvement in MDD (Cook et al, in press; Leuchter et al., 2005), the present study found a significantly greater week 1 decrease in MRF cordance in GMM responders as compared to non-responders. This finding provides support for a neurophysiologic basis underlying these symptom trajectory patterns. Subjects in the symptom volatility group showed an intermediate change in MRF cordance at week 1 that did not statistically separate from responders. One suggestion is that the cordance marker might simply act as a marker of clinical severity, or, of improvement regardless of the durability of response or the means by which it is attained. However, contrary to this interpretation, there was no significant association between the MRF marker and placebo response trajectories. Results of our analyses indicated that GMM outcome class differences in MRF cordance changes were moderated by time indicating that the assessment timepoint is important. Future studies should examine different timepoints and/or brain regions that could potentially help demarcate unique neurophysiologic features of subjects who express symptom volatility.

To our knowledge, this is the first report to relate symptom trajectory outcomes during treatment for MDD to a neurophysiologic marker. Findings should be interpreted within the limits of this study. First, this GMM approach should be replicated in an independent sample of subjects. Like any model, GMM is based on a set of assumptions (e.g., the assumption of normality within each latent class) that may not be completely fulfilled in any given data set. Second, the cause of any of the symptom trajectory patterns, including the symptom volatility pattern, cannot be determined from the present study. Response patterns in the various outcome classes may or may not be related to medication effects in varying degrees. For example, it is unknown to what extent fluctuating symptoms may be due to the natural course of illness and/or effects of treatment (e.g., medication side effects). Further, patients may have improvement in some symptoms while experiencing worsening of other symptoms; additional studies are needed to address changes in specific symptoms of interest. Third, subjects were treated with fluoxetine or venlafaxine; larger trials with random assignment to various antidepressants would be needed to assess potential effects of specific medications. Fourth, subjects in this study were not distinguished in terms of depressive subtypes (e.g., atypical depression) or comorbid anxiety disorders—clinical factors that can affect symptom ratings and SSRI response (Fava et al., 2008). Future studies of better-characterized subjects should be conducted to examine such clinical characteristics as potential moderators of response trajectory. Last, we examined QEEG cordance changes only at a single region and at a single timepoint previously shown to be related to neurophysiological or clinical effects of medication; it is possible that other groups of electrodes, other timepoints, or other EEG parameters might differentiate subjects who express a volatile symptom change outcome pattern.

Conflict of interest

Aimee M. Hunter, Ph.D. has no financial interests to disclose. Except for income received from her primary employer, Dr. Hunter has not received any financial support or compensation from any individual or corporate entity over the past three years for research

or professional service and there are no personal financial holdings that could be perceived as constituting a potential conflict of interest.

Bengt Muthén, Ph.D. is the co-developer of the Mplus software used in the GMM analyses. Except for income received from his software business, Dr. Muthén has not received any financial support or compensation from any individual or corporate entity over the past three years for research or professional service and there are no personal financial holdings that could be perceived as constituting a potential conflict of interest.

Ian A. Cook, M.D. has served as an advisor and consultant for Ascend Media, Bristol–Meyers Squibb, Cyberonics Inc., and Janssen. He has served on the Speaker's Bureau for Bristol–Meyers Squibb, Medical Education Speakers Network, Pfizer Pharmaceuticals Inc., and Wyeth Pharmaceuticals. Dr. Cook receives Research Support from Aspect Medical Systems, Cyberonics Inc., Eli Lilly & Company, Novartis Pharmaceuticals, Pfizer, Inc., and Sepracor.

Andrew F. Leuchter, M.D. has provided scientific consultation to or served on Advisory Boards for Aspect Medical Systems, Eli Lilly & Company, Novartis Pharmaceuticals, and MEDACorp. He serves on the speaker's bureau for Eli Lilly & Company, Wyeth Pharmaceuticals, and Pfizer Pharmaceuticals Inc. He has received research/grant support from the National Institute of Mental Health, Aspect Medical Systems, Eli Lilly & Company, MedAvante, Merck & Co., Novartis Pharmaceuticals, Pfizer, Inc., Wyeth Pharmaceuticals, and Vivometrics; and owns stock in Aspect Medical Systems (minor).

Funding source

This research of Aimee Hunter, Ian Cook, and Andrew Leuchter was supported by R01 AT002479 from the National Center for Complementary and Alternative Medicine. Grant support was also received from Aspect Medical Systems, Eli Lilly & Company and Wyeth. The research of Bengt Muthén was sponsored by the Center for Prevention and Early Intervention (P30 MH066247) and jointly funded by NIDA and NIMH, by grant no. MH40859 from NIDA and NIMH. None of the funding sources had a role in study design; in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the manuscript for publication.

Contributors

Aimee Hunter conceptualized the study, wrote the first draft of the manuscript, conducted literature searches, and statistical analyses. Bengt Muthén wrote Growth Mixture Modeling sections of the manuscript and conducted Growth Mixture Modeling analyses. Ian Cook designed primary studies and provided critical review of the clinical and EEG portions of the manuscript. Andrew Leuchter conceptualized the study, designed primary studies and provided critical review of the clinical and EEG portions of the manuscript. All authors contributed to and have approved the final manuscript.

Acknowledgements

We thank Xiao Chen and the UCLA Academic Technology Services (ATS) Statistical Consulting Group for statistical consultation, and we thank Jennifer Villalobos who assisted in preparation of the manuscript. We also thank Michelle Abrams R.N. and Barbara Siegman R.EEG.T., respectively, for clinical and EEG evaluations in the primary studies.

References

Akiskal HS. Validating 'hard' and 'soft' phenotypes within the bipolar spectrum: continuity or discontinuity? Journal of Affective Disorders 1993;73:1–5.

American Psychiatric Association. Practice guideline for the treatment of patients with bipolar disorder (revision). American Journal of Psychiatry 2002;159(Suppl.):1–50.

Bares M, Brunovsky M, Kopecek M, Stopkova P, Novak T, Kozeny J, et al. Changes in QEEG prefrontal cordance as a predictor of response to antidepressants in patients with treatment resistant depressive disorder: a pilot study. Journal of Psychiatric Research 2007;41:319–25.

Bares M, Brunovsky M, Kopecek M, Novak T, Stopkova P, Kozeny J, et al. Early reduction in prefrontal theta QEEG cordance value predicts response to venlafaxine treatment in patients with resistant depressive disorder. European Psychiatry 2008;23:350–5.

Beasley Jr CM, Koke SC, Nilsson ME, Gonzales JS. Adverse events and treatment discontinuations in clinical trials of fluoxetine in major depressive disorder: an updated meta-analysis. Clinical Therapeutics 2000;22:1319–30.

Chelben J, Strous RD, Lustig M, Baruch Y. Remission of SSRI-induced akathisia after switch to nefazodone. Journal of Clinical Psychiatry 2001;62:570–1.

Cook IA, O'Hara R, Uijtdehaage SH, Mandelkern M, Leuchter AF. Assessing the accuracy of topographic EEG mapping for determining local brain function. Electroencephalography and Clinical Neurophysiology 1998;107:408–14.

Cook IA, Leuchter AF, Morgan M, Witte E, Stubbeman WF, Abrams M, et al. Early changes in prefrontal activity characterize clinical responders to antidepressants. Neuropsychopharmacology 2002;27:120–31.

Cook IA, Hunter AM, Abrams M, Siegman B, Leuchter AF. Midline and right front brain function as a biomarker of remission in major depression. Psychiatic Research Neuroimaging, in press.

Cuijpers P, van Lier PA, van Straten A, Donker M. Examining differential effects of psychological treatment of depressive disorder: an application of trajectory analyses. Journal of Affective Disorders 2005;89:137–46.

Cusin C, Fava M, Amsterdam JD, Quitkin FM, Reimherr FW, Beasley Jr CM, et al. Early symptomatic worsening during treatment with fluoxetine in major depressive disorder: prevalence and implications. Journal of Clinical Psychiatry 2007;68:52–7.

Donovan SJ, Quitkin FM, Stewart JW, Ocepek-Welikson K, Harrison W, McGrath PJ, et al. Duration of antidepressant trials: clinical and research implications. Journal of Clinical Psychopharmacology 1994;14:64–6.

Fava M, Rush AJ, Alpert JE, Balasubramani GK, Wisniewski SR, Carmin CN, et al. Trivedi MH: difference in treatment outcome in outpatients with anxious versus nonanxious depression: a STAR D report. American Journal of Psychiatry 2008;165:342–51.

First MB, Spitzer RL, Gibbon M, Williams J. Structured clinical interview for DSM-IV axis I disorders, patient edition (SCID-I/P, version 2.0). New York; 1995.

Hamilton M. A rating scale for depression. Journal of Neurology, Neurosurgery & Psychiatry 1960;23:56–62.

Hunter AM, Cook IA, Leuchter AF. The promise of the quantitative electroencephalogram as a predictor of antidepressant treatment outcomes in major depressive disorder. Psychiatric Clinics of North America 2007;30:105–24.

Kaplan EM. Antidepressant noncompliance as a factor in the discontinuation syndrome. Journal of Clinical Psychiatry 1997;58(Suppl. 7):31–5.

Leuchter AF, Uijtdehaage SH, Cook IA, O'Hara R, Mandelkern M. Relationship between brain electrical activity and cortical perfusion in normal subjects. Psychiatry Research 1999;90:125–40.

Leuchter AF, Cook IA, Hunter AM, Morgan ML, McGrouther CC, Abrams M. Functional brain changes associated with remission of major depressive disorder. Poster presented at the Society of Biological Psychiatry Meeting, Atlanta, GA., 2005.

Leuchter AF, Cook IA, DeBrota DJ, Hunter AM, Potter WZ, McGrouther CC, et al. Changes in brain function during administration of venlafaxine or placebo to normal subjects. Clinical EEG and Neuroscience 2008;39:175–81.

Muthén B, Muthén LK. Integrating person-centered and variable-centered analyses: growth mixture modeling with latent trajectory classes. Alcoholism: Clinical and Experimental Research 2000;24:882–91.

Muthén B, Shedden K. Finite mixture modeling with mixture outcomes using the EM algorithm. Biometrics 1999;55:463–9.

Muthén B, Muthén L. Mplus version 5 software users guide;1998–2008.

Muthén B, Brown CH, Masyn K, Jo B, Khoo ST, Yang CC, et al. General growth mixture modeling for randomized preventive interventions. Biostatistics 2002;3:459–75.

Muthén B. Latent variable analysis: growth mixture modeling and related techniques for longitudinal data. In: Kaplan D, editor. Handbook of quantitative methodology for the social sciences. Newbury Park, CA: Sage Publications; 2004. p. 345–68.

Muthén B, Asparouhov T. Growth mixture modeling: analysis with non-Gaussian random effects. In: Fitzmaurice G, Davidian M, Verbeke G, Molenberghs G, editors. Advances in longitudinal data analysis. Boca Raton: Advance; 2008. p. 143–65.

Muthén B, Brown H. Estimating drug effects in the presence of placebo response: Causal inference using growth mixture modeling. Statistics in Medicine, in press

Muthén B, Brown H, Leuchter AF, Hunter AM. Applying growth mixture modeling to randomized trials of depression medication. In: Strout PE, editor. Causality and psychopathology: finding the determinants of disorders and their cures. Unpublished technical report 2008. Available from: http://www.statmodel.com/download/APPA2008March27.pdf.

Nierenberg AA, Farabaugh AH, Alpert JE, Gordon J, Worthington JJ, Rosenbaum JF, et al. Timing of onset of antidepressant response with fluoxetine treatment. American Journal of Psychiatry 2000;157:1423–8.

Nylund KL, Asparouhov T, Muthén BO. Deciding on the number of classes in latent class analysis and growth mixture modeling: a Monte Carlo simulation study. Structural Equation Modeling: A Multidisciplinary Journal 2007;14: 535–69.

Papakostas GI, Petersen T, Sklarsky KG, Nierenberg AA, Alpert JE, Fava M. Timing of clinical improvement and symptom resolution in the treatment of major depressive disorder. Psychiatry Research 2007;149:195–200.

Perahia DG, Pritchett YL, Kajdasz DK, Bauer M, Jain R, Russell JM, et al. A randomized, double-blind comparison of duloxetine and venlafaxine in the treatment of patients with major depressive disorder. Journal of Psychiatric Research 2008;42:22–34.

Posternak MA, Zimmerman M. Is there a delay in the antidepressant effect? A meta-analysis. Journal of Clinical Psychiatry 2005;66:148–58.

Quitkin FM, Rabkin JG, Ross D, Stewart JW. Identification of true drug response to antidepressants. Use of pattern analysis. Archives of General Psychiatry 1984;41:782–6.

Quitkin FM, Rabkin JD, Markowitz JM, Stewart JW, McGrath PJ, Harrison W. Use of pattern analysis to identify true drug response. A replication. Archives of General Psychiatry 1987;44:259–64.

Quitkin FM, McGrath PJ, Rabkin JG, Stewart JW, Harrison W, Ross DC, et al. Different types of placebo response in patients receiving antidepressants. American Journal of Psychiatry 1991;148:197–203.

Quitkin FM, McGrath PJ, Stewart JW, Taylor BP, Klein DF. Can the effects of antidepressants be observed in the first two weeks of treatment? Neuropsychopharmacology 1996;15:390–4.

Ramaswamy V, DeSarbo W, Reibstein D, Robinson W. An empirical pooling approach for estimating marketing mix elasticities with PIMS data. Marketing Science 1993;12:103–24.

Sachs GS, Koslow CL, Ghaemi SN. The treatment of bipolar depression. Bipolar Disorders 2000;2:256–60.

Stassen HH, Angst J, Hell D, Scharfetter C, Szegedi A. Is there a common resilience mechanism underlying antidepressant drug response? Evidence from 2848 patients. Journal of Clinical Psychiatry 2007;68:1195–205.

Stulz N, Lutz W. Multidimensional patterns of change in outpatient psychotherapy: the phase model revisited. Journal of Clinical Psychology 2007;63:817–33.

Stulz N, Lutz W, Leach C, Lucock M, Barkham M. Shapes of early change in psychotherapy under routine outpatient conditions. Journal of Consulting and Clinical Psychology 2007;75:864–74.

Trivedi MH, Rush AJ, Wisniewski SR, Nierenberg AA, Warden D, Ritz L, et al. STAR*D study team. Evaluation of outcomes with citalopram for depression using measurement-based care in STAR D: implications for clinical practice. American Journal of Psychiatry 2006;163:28–40.