Abstract and Relevance

The timely selection of the best treatment for patients with depression is critical to the goal of improving remission rates. Due to the biological heterogeneity and variable symptom presentation of depression, it is unlikely that a single clinical or biological marker can guide treatment selection. Rather, a biosignature developed from a systematic exploration of a group of several clinical and biological markers is more likely to be successful. Two types of biosignatures are needed to achieve improved outcomes: 1) biosignature to maximize the selection of optimal treatment for individual patients at beginning of treatment (moderators) and 2) biosignatures to identify indicators of eventual outcomes early in treatment (mediators). This approach has great potential to personalize treatment and to begin to characterize the biology of treatment response.

We propose a randomized, placebo-controlled trial of a serotonin selective reuptake inhibitor citalopram (CIT) and placebo (PBO) for participants with major depressive disorder (MDD) in which we will assess a comprehensive array of carefully selected clinical (e.g., anxious depression, early life trauma, gender) and biological (i.e. neuroimaging, electrophysiology and behavioral neuropsychiatric) moderators and mediators of outcome. Using innovative statistical approaches the identified moderators and mediators will then be used to develop a differential depression treatment response index (DTRI).

The proposed study is a randomized two-stage trial design with 400 MDD patients randomly assigned to one of two treatments under masked conditions (CIT vs. PBO; n=200 each). At the end of Stage 1, nonresponders to 8 weeks of CIT will be switched to bupropion (BUP), and nonresponders to PBO will be switched to CIT. This two-stage approach is similar to a Sequential Multiple Assignment Randomized Trial (SMART) design, which allows for the exploration of multiple treatments in individual patients

This study will be conducted under the joint leadership from Columbia University and University of Texas Southwestern Medical Center together with Massachusetts General Hospital, University of Michigan, University of Pittsburgh and McLean Hospital. This study brings together researchers with extensive experience in conducting large clinical trials together with experts at the forefront of the neurobiology of depression, including: clinical trials (Trivedi, Weissman, McGrath, and Fava), neuroimaging (Phillips, Parsey, and Buckner), neurophysiology (Bruder and Pizzagalli), clinical predictors (Weissman, Trivedi, McGrath, Fava, Kurian, Morris, and Oquendo). This team will also be guided by a highly qualified group of biostatisticians (Petkova, Kraemer, Ogden, and Carmody), with specific expertise in emerging statistical methods to develop disease biomarkers using complex biomarker data.

This study will examine multiple carefully selected clinical and biological markers, using both existing state-of-the-art technologies as well as pioneering, innovative approaches. Evaluation of the usefulness of these markers in a carefully conducted clinical trial comparing an antidepressant to placebo will assist in developing a depression treatment response index (DTRI) to help clinicians match treatments to patients with MDD, resulting in timely selection of treatments best suited for individual patients and thus approaching personalized treatment.

1. INTRODUCTION

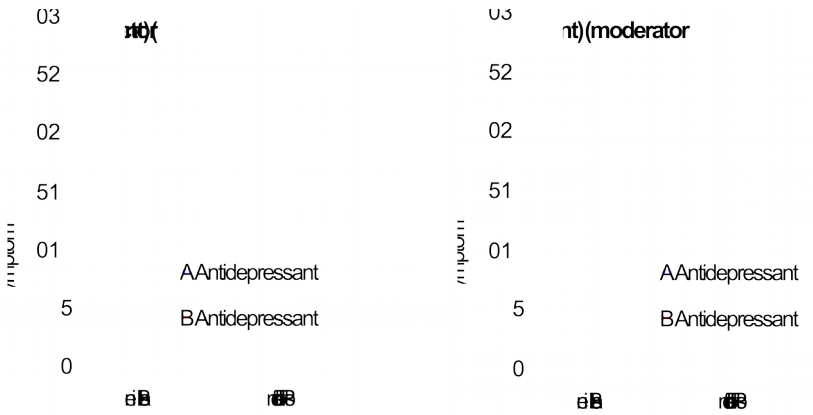
1.1.General Approach in Conducting the Research: **EMBARC** (**E**stablishing **M**oderators/Mediators for a **B**iosignature of **A**ntidepressant **R**esponse in **C**linical care) uses a two-stage trial (Stage 1: 8 weeks; Stage 2: 8 weeks) design with 400 MDD patients diagnosed via SCID/P and randomized to two treatment conditions: citalopram (CIT) vs. placebo (PBO), n=200 each in Stage 1. The design will identify Stage 1 nonresponders, and assign them to a new active treatment. The two-step approach is similar to a Sequential Multiple Assignment Randomized Trial (SMART) design,[1] which has been proposed as a way to develop sequences of treatments, or adaptive treatment strategies that are guided by a patient's responses to prior treatments. Specifically, participants not responding to CIT will be switched to BUP, and nonresponders to PBO will be switched to CIT for an additional 8 weeks. Responders to PBO and CIT will remain on their respective treatment for an additional 8 weeks. In view of the relatively high rate of placebo response commonly seen with MDD, this protocol will evaluate moderators and mediators via a double-blind, randomized, placebo-controlled trial of an SSRI. Selective serotonin reuptake inhibitors (SSRIs) represent the current first-step standard of care for MDD. Citalopram is a commonly used, generically available SSRI which offers few pharmacokinetic interactions and is therefore a treatment of choice in some patient subgroups. In order to evaluate moderators and mediators of an alternative nonserotonergic antidepressant, nonresponders to citalopram will be switched to bupropion. This switch arm will also provide clues to establishing a biosignature for nonresponse (for those not responding to both CIT and BUP in the two-step sequence). Placebo nonresponders from Stage 1 given CIT in Stage 2 serve as an enriched sample to confirm Stage 1 moderators and mediators to citalopram treatment response. The 8-week extension for responders to citalopram and placebo will also provide a distinct opportunity to assess the biomarker predictors of stable response over 16 weeks.

The array of potential moderators and mediators was developed based on a critical review of the literature as well as the direct work of our multidisciplinary team. This team has the expertise and track record to: a) successfully conduct the clinical trial within the time frame allotted; b) obtain and analyze high quality biomarkers; c) collect, store, and manage the complex clinical and biological data from multiple sites; and d) use advanced data analytic tools to integrate the markers and develop an innovative differential depression treatment response index (DTRI) for personalized treatment. The DTRI is a biosignature index designed to guide individual treatment decisions. Such an index is a critical first step in developing an index with the potential to have a large impact on clinical practice. The placebo controlled design allows us to identify specific clinical and biological markers of an active treatment. This study is the first step in the development of biosignatures of treatment, and may also advance the understanding of the neurobiology of depression.

In addition to the proposed measures, the investigators will collect samples of DNA, mRNA, and plasma proteins, selected by expert consultants in the fields of molecular genetics, transcriptomics, and proteomics, both before and during treatment, and deposit these samples in a repository for future study by other investigators to explore other innovative approaches to the development of biosignatures of treatment response.

Importantly, this initial study will develop standards for the collection, processing, and evaluation of clinical and biological markers. This systematic exploration of a wide array of clinical and biological markers for well-defined phenotypes will serve as a resource for the field, and establish the infrastructure to conduct future, studies.

Figure 1.



We propose an approach to the use of clinical and biological markers as moderators and mediators as defined by Kraemer et al.[2] Moderators are pre-treatment clinical or biological characteristics that are associated with differential treatment outcomes for different treatments. There is often confusion between a "predictor" and a "moderator." A predictor of treatment outcome is a variable that is correlated with subsequent outcome irrespective of type of treatment(s). ***Moderators*** on the other hand are correlated with one treatment but not another. As presented in Figure 1, a predictor is a

variable that predicts response for both treatments, while in contrast a moderator predicts differential response for the two treatments. Therefore, a moderator provides useful information in favor of selecting one treatment over another and can only be identified in studies with multiple treatments.

In contrast to moderators, ***mediators*** are defined as changes early in treatment (after one or 2 weeks) and represent early indicators of eventual treatment outcome. Examples of potential mediators include quantitative electroencephalogram (qEEG) recordings, measurements of brain function via functional MRI (fMRI) activity, changes in behavioral performance, and changes in the transcriptome or in plasma signaling proteins. Mediators inform clinicians about when to stop an ineffective treatment early, and may provide insight into the underlying pathophysiology of treatment response, and, in turn, guide alternative treatment approaches for specific individuals.

Three approaches were used by our team to identify potential clinical and biological candidates, prioritizing the use established markers that have been shown to be associated with treatment outcome, though whether as predictors or as moderators has not been adequately investigated. Clinical markers were identified based on best available evidence from the literature, including results from the NIMH-sponsored, multicenter Sequenced Treatment Alternatives to Relieve Depression (STAR*D) study,[3-5] with a specific focus on differential treatment outcomes. We carried out a systematic review of the literature on clinical and biological markers of antidepressant response in MDD, including potential neurobiological moderators and mediators based on the best current models of depression.[6-10] Finally, we have included biological measures for which we have unpublished preliminary data. Through this process, we determined that the most promising clinical and biological marker candidates to assess are:

1. Clinical Moderators – anxious depression, early life trauma, gender, melancholic and atypical depression, anger attacks, Axis II disorder, hypersomnia/fatigue, chronicity of depression, and

2. Neuroimaging Moderators – fMRI measures of activity and functional (effective) connectivity to two tasks (an implicit emotion processing and regulation task assessing amygdala-anterior cingulate gyrus [ACC] circuitry, and a reward processing task assessing ventral striatal-orbitomedial prefrontal cortex [OMPFC] circuitry), resting state intracerebral connectivity, pulsed arterial spin labeling (PASL; a measure of regional cerebral blood flow), diffusion tensor imaging, and structural MRI measures of cortical thickness.

3. Electrophysiology Moderators – Pretreatment alpha and theta EEG power, and source localization measures of theta activity in the rostral ACC (qEEG) and Loudness dependency of auditory evoked potentials (LDAEP).

4. Behavioral Phenotyping Moderators – psychomotor slowing (as measured through speeded RT and word fluency tasks), cognitive control (as measured by interference and post-error adjustments), working memory, and reward responsiveness (as measured through a probabilistic reward task).

5. Neuroimaging/ Electrophysiology Mediators − changes in the above fMRI measures, changes in resting state connectivity, changes in PASL, and changes in the above EEG measures from baseline to one week after treatment initiation.

We plan to analyze this broad array of moderator and mediator biomarkers individually and then assess how these markers are related to each other with innovative statistical methods within a systematic theoretical framework that will be developed through the work of this group. The findings will then be integrated to form the **differential depression treatment response index (DTRI)**. Additionally, biomaterials such as plasma, mRNA, and DNA will be collected to explore, and later test when funding becomes available, the predictive value of these biomarkers as moderators and mediators. Since it is critical to minimize placebo response in clinical trials, participants in this study will be required to have a history of chronic (lasting over two years) MDD and/or early onset recurrent MDD.

1.2. Mechanisms and Treatment of MDD

1.2.1. Rationale for Inclusion of Placebo (Stage 1, CIT vs. PBO): The inclusion of a placebo control group provides the distinct opportunity to differentiate between nonspecific changes that are due to environmental or psychological factors versus those that are the result of treatment. Furthermore, including a PBO assists in identifying changes, which occur in the earlier stages of treatment that mediate treatment response. It will also allow for the identification of baseline predictors of a treatment response, although identification of moderators or mediators that are specific to a treatment requires that an alternative active treatment be part of the design. In addition, it may allow for the identification of markers that indicate that recovery from the current depressive episode is likely without medication treatment.

1.2.2. Selecting Distinct Treatments to Optimize Stage 2 Outcomes (CIT and BUP): An essential step in developing the most clinically useful biosignature entails comparing established treatments to test for moderating effects, which we propose to do among Stage 1 nonresponders. Our treatment selection process for Stage 1 nonresponders is guided by two primary tenets: 1) select a putatively distinct treatment with a theorized biological mechanism that relates to our proposed moderators, and 2) choose clinically relevant treatments that have been rigorously tested through randomized controlled trials in real-world settings. Our resulting selections, a serotonergic antidepressant, citalopram (CIT) and a nonserotonergic antidepressant, bupropion (BUP), represent the most likely treatments for which a biosignature would have a large impact on the current treatment selection process in real-world practices. In addition, the inclusion of CIT for placebo nonresponders enhances the capacity to identify a biosignature for citalopram response.

1.2.2a. Brief Review of the Underlying Mechanisms of Action for the Proposed Treatments

CIT: Based on a recent analysis, treatment with antidepressant medications has steadily increased over the past decade.[11] Of these, the most commonly prescribed agents are the SSRIs, particularly since all but one of the SSRIs are available in generic formulations. Based on results from STAR*D and other large effectiveness trials, the SSRIs and other newer antidepressant agents are modestly successful at achieving remission, and are mechanistically designed to treat the monoaminergic model of MDD.

The immediate action of the SSRIs is to inhibit the neuronal reuptake of serotonin by blocking the serotonin transporter.[12, 13] The SSRI paroxetine appears to have mild noradrenergic reuptake inhibition at high doses, [12, 13] while the SSRI fluoxetine, particularly its R stereo-isomer, has mild 5HT2A and 5HT2C antagonist activity, which may explain the increase in norepinephrine and dopamine in the prefrontal cortex of animals treated with fluoxetine,[14] as well as mild noradrenergic reuptake inhibition.[12, 14] The SSRIs have minimal or no affinity for cholinergic receptors, with the exception of paroxetine which is a weak cholinergic receptor antagonist.[12, 14] The effects of SSRIs on various histaminergic and α-adrenergic receptors are negligible.[12, 14] CIT and escitalopram are probably the most selective SSRIs from a pharmacological standpoint, and the least likely to be involved in the modulation of other neurotransmitter systems. Hence we selected CIT, which is available in a generic formulation (escitalopram is not) and was used in Level 1 of STAR*D, as one of our treatment arms.

PBO: A placebo control will be used to establish the frequency and magnitude of changes in clinical and biological marker endpoints that may occur in the absence of pharmacologically active treatment. Thus, the inclusion of a PBO group will provide valuable information to characterize treatment specific biological and clinical markers. Lastly, PBO nonresponders provide an enriched sample to further explore biomarkers of CIT response in a group of participants whose PBO response is likely less, and thus is biologically more homogenous.

BUP: The mechanism of action of bupropion (BUP) is unclear but has been postulated to be primarily related to the inhibition of the reuptake of both dopamine and norepinephrine, but it seems fair to characterize it as nonserotonergic and thus distinct from CIT.[15] BUP and its metabolites have shown to be able to inhibit striatal uptake of the selective DAT-binding radioligand (11)C-beta CIT-FE in vivo achieving DAT occupancy ranging from approximately 14%[16] to 26%[17, 18] at therapeutic doses. BUP has also been reported to have mild affinity for the norepinephrine transporter, although some researchers have argued that the effect of BUP on norepinephrine is primarily through an increase in presynaptic norepinephrine release.[19] Regardless of the exact mechanism, the overall effect of BUP appears to be a dose-dependent increase in brain extra cellular dopamine and norepinephrine concentrations.[19-21] In addition, BUP also appears to noncompetitively inhibit the α3β2, α3β4-, and α4β2- nicotinic acetyl cholinergic receptors in vitro.[22] At least four stereo isomers of two major metabolites of BUP (S,S-, R,R hydroxybupropion, R,R-threohydrobupropion) were also reported to noncompetitively inhibit α3β4- nicotinic acetyl cholinergic receptors in vitro, although not as potently as BUP.[22] Because its mechanism of action does not involve serotonin but does appear to involve catecholamines, we have selected BUP as the comparative active treatment arm in Stage 2.

1.3. Rationale for Proposed Approach:

1.3.1. Shortcomings of Current Antidepressants: The largest clinical trial of MDD ever conducted, STAR*D, indicates that 2/3 of patients treated with a first step antidepressant do not achieve remission of symptoms.[5] Furthermore, successive treatment steps lead to diminishing remission rates.[23] Additionally, a large number of patients discontinue treatment prematurely due to side effects.[5] The trial and error method currently used in clinical practice often leads to repeated failures before an effective treatment is identified.

1.3.2. Biosignatures to Individualize Depression Treatment: Given this relative ineffectiveness of treatments for depression and resulting practice of trial and error multiple treatment step algorithms, there is an urgent need to identify factors that can be used to personalize treatment (i.e., markers that maximize effectiveness and minimize the risk for toxicity). Similar efforts have been successfully used in treating other medical illnesses.[24-29] The National Institute of Mental Health's (NIMH) RFA-MH-10-040 recognizes the need to develop personalized treatments for individuals with MDD and "discover panels of promising potential biomarkers (i.e., biosignatures) that are predictive of treatment outcomes." It is likely that several clinical and biological markers from several domains that we have identified as candidates will define biosignatures that are predictive of treatment outcome for individual patients with MDD. We will primarily employ strategies to combine biomarkers and pursue the most promising markers through this effort.

1.3.3. The Challenge in Developing a Biosignature for MDD: The development of biomarker predictors of antidepressant response languished after multiple candidates, most notably the dexamethasone suppression test (DST), proved to have inadequate prognostic clinical utility[30] However, the emergence of new technologies, including neuroimaging and neurophysiology, has sparked new interest in developing biomarkers that might predict antidepressant response.

Because of limited understanding of the pathophysiology of MDD and the limited range of the mechanism of action of available antidepressants (monoaminergic uptake inhibition or receptor modulators), we are currently unable to match treatments to patients. Subgroups of patients with MDD who share the same set of biomarker measures could respond to drugs with similar mechanisms of action. We designed the current study to provide the data so that we can develop a formula that weights the importance on biomarker findings. We foresee that for us to select the optimal biomarkers that need to be administered to develop personalized treatment, we will have to take into account interactions between these measures and their relationship to clinical features. Thus, it is likely that a small set of tests will ultimately suffice to match the goal of guiding treatment selection. The result will be for clinicians to make the right decision for differential therapy that is not based on intuition and impressions, but instead on a more precise combination of clinical features and objective bioanalytical measures.

1.3.4. Shortcomings of Previous Treatment Prediction Trials: Previous research has focused on predictor variables and provided evidence of markers associated with poor outcome. However, the clinically relevant question of how to match patients with specific treatments using relevant behavioral and biological markers has not been well studied. To date, only a few small exploratory studies were conducted, and were primarily designed to examine single biomarkers. To advance the field, it is necessary to identify a panel of moderator variables that are likely to guide clinicians' treatment decisions for patients with MDD.

1.3.4.a. Prior Clinical Predictor Trials: Typically, depression studies identify predictors that explain only a small percentage of the systematic variance, often less than 10%.[31-33] STAR*D provided clinical information about treatment response throughout a number of treatment levels. In general, for level 1 treatment with CIT, Trivedi and colleagues found that comorbid psychiatric and/or medical conditions, greater initial depressive symptom severity, and lower socioeconomic status (SES) predicted a lower likelihood for achieving remission.[5] Nonresponders to CIT were then randomized to various different pharmacological and psychological treatments in level 2. Rush and colleagues assessed clinical moderators of treatment response among these level 2 treatments. Unfortunately, when examined, no significant clinical moderators (i.e., variables that differentially predicted treatment response to an individual medication) were identified.[34]

STAR*D also assessed some specific phenotypes that were associated with treatment response. For example, despite the fact that the DSM-IV-TR does not currently include a specifier for depression with anxious features, a number of studies have described this common correlate in the literature.[35, 36] Based on research from STAR*D, 46% of patients met baseline criteria for anxious depression.[35] Patients with MDD and comorbid anxiety are more likely to suffer increased disease burden and longer duration of illness.[35, 36] Furthermore, most studies, including STAR*D, have shown that patients suffering from anxious depression are less likely to respond and remit following antidepressant treatment than are those with nonanxious depression.[37]

1.3.4.b. Prior Biological Predictors: Biological predictors of antidepressant response originally revolved around DST, measures of monoamines and their metabolites in urine, cerebrospinal fluid, or blood.[38] However, these did not prove to be sufficiently predictive of outcome.

1.3.4.c. Imaging Predictors/Moderators: Structural and functional magnetic resonance imaging (fMRI) may be used to help predict treatment response. The poorer responses to treatment observed in patients with comorbid anxiety may be understood through application of recent fMRI paradigms to assess key amygdala-ACC circuitry implicated in implicit emotion processing and regulation. Our collaborative group[39, 40] has reported that patients with generalized anxiety disorder (GAD) differ significantly from healthy controls in patterns of ACC activity on tasks of implicit emotion processing and regulation, and that patterns of ACC activity to such tasks in patients with MDD predict future outcome to antidepressant treatment. As described below, patients with MDD comorbid with GAD are similar to GAD patients without comorbid depression on these tasks and differ from major depressives without GAD. Employment of an implicit emotion processing and regulation tasks with fMRI is therefore one method for assessing ACC function, that relates to anxiety in major depression, and that could be useful pre-treatment signature for predicting and moderating treatment response in MDD. Our collaborative group has also reported reduced ventral striatal activity to reward in patients with MDD relative to healthy controls[41] that may in turn be associated with anhedonia in MDD. Employment of a reward task with fMRI is therefore a method for assessing ventral striatal – and reward circuitry - function, that in turn has potential to be a useful pre-treatment signature for predicting and moderating treatment response in MDD. We will also employ resting state connectivity, pulsed arterial spin labeling (PASL), and diffusion tensor imaging (DTI) to obtain measures of low frequency BOLD fluctuations (LFBF) at rest, regional cerebral blood flow, and white matter integrity, respectively, in our key emotion processing, regulation, and reward neural circuitries of interest. In addition, structural MRI will measures cortical thickness, that may also be useful pre-treatment signatures predicting and/or moderating treatment response in MDD.

1.3.4.d. Electrophysiology Predictors/Moderators: We aim to assess pretreatment EEG measures of brain activity to differentially predict antidepressant response. Pretreatment alpha and theta power have been shown to be associated with antidepressant response.[42-46] Furthermore, compelling evidence finds a significant association between increased pretreatment resting theta activity in the rostral ACC and antidepressant response (including SSRIs).[47-49]

In addition, findings suggest that loudness dependence of auditory evoked potential (LDAEP) correlates with serotonergic treatment response. Specifically, prior studies have shown that patients with higher pretreatment loudness dependence – assumed to reflect blunted sertonergic activity – responded well to SSRI, while responders to the selective noradrenergic reuptake inihibitor, reboxetine, had reduced pretreatment loudness dependence.[45, 50-52] Similarly, response to the SSRI citalopram was associated with strong LDAEP, while response to reboxetine was associated with weak LDAEP.[50]

1.3.4.e. Behavioral Predictors/Moderators: We will measure reaction time (interference and post-error adjustments) during performance of the implicit emotion processing and regulation and reward processing neuroimaging tasks as key behavioral measures associated with neuroimaging measures of ACC- and ventral striatal-centered neural circuitry that in turn have potential to be useful pre-treatment signatures predicting and/or moderating treatment response. In addition, we will measure pre-treatment psychomotor slowing, cognitive control (particularly post-error behavioral adjustments), working memory performance, and reward responsiveness, which have preliminary evidence that they may be able to differentially predict treatment response.

1.3.4.f. Summary of Prior Predictor Literature: Past clinical and biological predictors have not been useful for treatment selection. Two key factors have led to this problem: first, it is clear that depression is a heterogenous complex disorder that involves multiple systems and neural circuits. Secondly, treatment response is highly variable. Past research has been limited and has examined how a single marker affects treatment outcome to a single treatment at any given time (i.e., prediction). Given these limitations, it is not surprising that there are no readily available markers to assist treatment selection for patients. This study aims to provide an initial step for the identification of clinical and biological markers and their interactions. It will also develop methods and procedures in which this approach can be applied to a wider array of treatments for MDD. Clinical decisions based on such multiple independent indices should be stronger than those based on any single moderator alone.

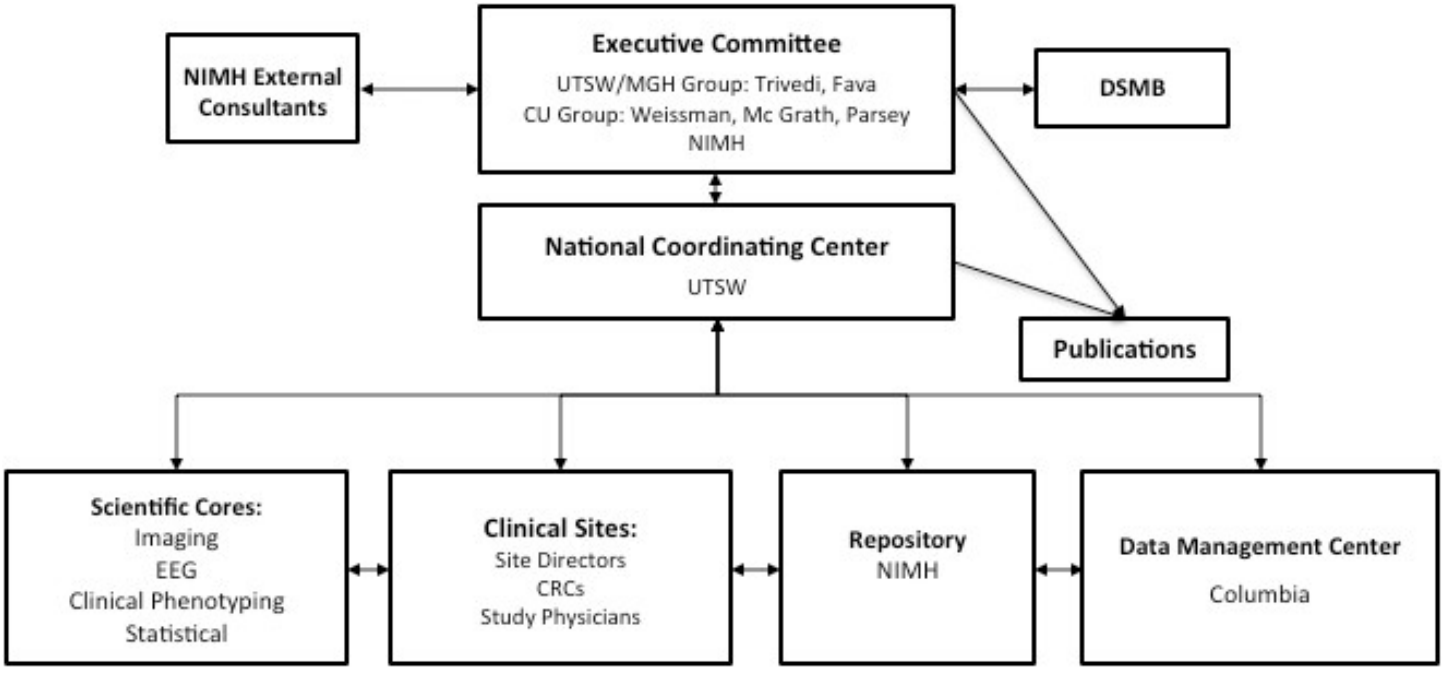
1.4. A Strategy for Developing a Biosignature Model: In this study we have taken an **efficient neuroscience systems approach** to studying clinical and biological markers.[53] We began with the central organizing hypothesis that the neurobiology of depression and antidepressant response is expressed at multiple hierarchical levels of systems organization. The levels of organization we focused on were neural (e.g., fMRI, EEG), behavioral (e.g., clinical/cognitive tasks), developmental (e.g., trauma history, history of prior treatment response), and social/vocational (e.g., quality of life measures). Because of the hierarchical nature of the interactions across these levels of organization, it is possible that a higher order marker (social function) would reflect a process occurring at a more basic level of organization as well as emergent properties unique

to that level. Thus, we do not consider the multidimensional biomarker approach to introduce substantial redundancy. Similarly, in those cases where more than one biomarker is evaluated within a particular level of systems organization (as in the evaluation of neural systems using EEG and fMRI), measures are included because they are thought to reflect orthogonal dimensions of homeostatic/allostatic processes at that level of organization.

We view this study as part of a larger effort to accelerate the development of personalized medicine research in depression. Nearly all of the biomarkers employed in this study are validated by multiple single-site trials or a single multicenter study (as in STAR*D). Thus, this study represents a "bridge" between single site or smaller studies designed to "nominate" candidate predictive markers and definitive trials designed to test the impact of biomarkers supported by definitive studies on treatments that have been similarly validated. Because it combines features of hypothesis-generating and hypothesis-testing, we propose to test the selected biomarkers individually and then to build multivariate models that are supported by univariate associations.

As indicated above, early studies in biological predictors of response in depression revolved around measuring metabolites of monoamines based on theories of synaptic monoamine dysfunction with less focus on brain circuitry and anatomical substructures.[38] Over the past decade, the development of a number of structural and functional imaging tools, as well as molecular mapping techniques, has fostered development of approaches that more commonly incorporate brain circuits and anatomy to further our understanding of why some individuals become depressed. However, less progress has been made in employing these techniques to predict response to treatment. Understandably then, there are important interactions among various biological markers (particularly neuroimaging results) as well as other variables such as environment, specific demographics (e.g., gender) etc. While strides toward finding a single marker to provide a great deal of information regarding response to one or another treatment have been made, any one biomarker has limited predictive power. It is more likely that we will ultimately observe interactions among variables on a clinical characteristic such as treatment response. Thus, key interactions across domains are likely to be important in understanding depression and predicting response. This study will emphasize fMRI, and qEEG as key biological moderators of response and will compare and link them to clinical and behavioral moderators.

1.5. Multidisciplinary Approach to Evaluating Clinical/Biological Markers and the Conduct of a Multisite Clinical Trial: This proposal brings researchers with extensive experience in clinical trials (Drs. Trivedi, McGrath, Fava, Weissman), neuroimaging (Drs. Phillips, Parsey, Buckner), quantitative electroencephalography (qEEG) (Drs. Bruder, Pizzagalli, and Iosifescu), clinical predictors (Drs. Fava, McGrath, Trivedi, Kurian, Morris, and Oquendo), and behavioral phenotypes (Drs. Pizzagalli, Bruder, and Phillips). This team will also be guided by a highly qualified group of biostatisticians (Petkova, Kraemer, Ogden, and Carmody) who will provide both the day-to-day oversight for data and statistics coordination as well as guidance on methodological innovations essential to this proposal.



1.5.1. Approach to Clinical Trial Project Management: Inter-site Collaboration, Coordination, and Quality Assurance: This study will be conducted using the experience gained over the last 10 years in managing multiple sites and collecting a wide array of data (e.g., STAR*D, Combining Medications to Enhance Depression Outcomes [CO-MED] and studies at CU). The study scientific team includes experts in clinical trials, clinical phenotyping as well as biological markers. They will oversee the collection, quality control, transmission, and analysis of clinical, neuroimaging, qEEG, and behavioral phenotypes. They will be assisted by other collaborators noted above. The experts overseeing biological marker collection and analysis have a long history of successful completion of these tasks and their functions will be seamlessly integrated into the existing infrastructure through organized Core Teams (see Sections 4.1, 4.2. and 5).

Coordination across all clinical trial components of the study and monitoring are provided by the National Coordinating Center (NCC) in Dallas (Madhukar Trivedi). Coordination of the data management and data analyses of the study and monitoring will provided by the Data Management Center (DMC) in New York (Myrna Weissman). The NCC and the DMC will coordinate study startup and implementation, publications, and dissemination of study findings. The statistical experts will also provide guidance on innovative analytic approaches so essential to DTRI development.

Each of four Clinical Sites (CSs) (UT Southwestern Medical Center [UTSW], Columbia University [CU], Massachusetts General Hospital [MGH], and University of Michigan [UM]) will screen, assess, and randomize subjects, provide psychopharmacology treatment, collect clinical/biological marker data, and follow the same protocol overseen by the NCC. The leaders of each CS (Drs. Trivedi, Kurian, Weissman, McGrath, Parsey, Fava, McInnis) have demonstrated the capacity to recruit, enroll, treat and retain depressed patients, and have extensive experience in the collaboration and implementation of complex multi-center studies.

The organizational structure builds on the core committee/administrative structure used effectively in STAR*D and CO-MED (STAR*D enrolled more than its planned 4000 participants and CO-MED completed enrollment before its aggressive targeted enrollment completion date). We have generated over 100 publications of findings and methods from STAR*D and related studies. Committees include the Steering Committee, NCC Study Management Team, Operations Committee, and Publications Committee. For this study we will also add four Scientific Cores: neuroimaging (led by Drs Phillips, Parsey, and Buckner), qEEG (led by Drs. Bruder and Pizzagalli), clinical predictors (Drs. McGrath, Fava, Trivedi, Weissman), and behavioral phenotypes (led by Drs. Pizzagalli, Bruder, and Phillips).

1.5.2. The Role of the National Coordinating Center (NCC): Dr. Trivedi will lead the NCC, which will be responsible for the operational leadership of the trial. The primary purpose of the NCC will be to ensure the execution of the study in accordance with the protocol by providing central coordination across all components of a large and geographically broad infrastructure. The NCC will: promote optimal performance in the start-up and implementation phases of the protocol, oversee the data collection and analyses, and coordinate the publication and dissemination of study findings.

To provide scientific leadership, direction, and oversight for the study in conjunction with collaborators at CU. This aim is accomplished by developing the protocol for the study, providing overall scientific management and oversight of the execution of the clinical trial; maintaining scientific integrity; coordinating the implementation of the protocol with scientific investigators at the DMC and RCs; providing ongoing review for possible protocol changes; providing scientific oversight and input for the development, review, and approval of ancillary studies for submission to NIMH; and developing and participating in planning and writing manuscripts to ensure productivity and dissemination of scientific findings through high quality peer-reviewed publications. To this end, the NCC will be responsible for organizing a publications committee, which will oversee the publication and dissemination of study findings.

To provide clinical oversight, training, and quality control to ensure rapid and scientifically rigorous implementation of the protocol. This aim is accomplished through finalizing assessment measures and clinical and data procedures; developing the Clinician and Clinical Research Coordinators (CRC) Procedures Manual and training procedures for Study Clinicians, Regional Center Directors (RCDs), Clinical Site Directors (CSDs), and CRCs; providing training and oversight of clinical and research study staff; providing quality control for clinical medication management and data management procedures, ensuring ongoing participant safety – including adverse and serious adverse event (AE/SAE) monitoring/reporting and providing this information to appropriate entities (i.e., DSMB). Finally, the NCC will oversee changes in study clinical procedures when indicated for scientific or safety issues.

The Coordinating Center will provide rater training and oversee annual inter-rater reliability assessments under the supervision of Drs. Morris and Trivedi

To provide administrative oversight and management of the study to ensure successful implementation of the protocol. This is accomplished by maintaining contracts with study cites; ensuring that study services meet the needs of a diverse participant population; coordinating communications, including arranging teleconferences, face-to-face meetings, and site visits; and maintaining communication logs of meetings, teleconferences, emails, and minutes of committee meetings. The NCC ensures that all required reports are submitted to NIMH and communications are maintained with the DMC and the Data Safety Management Board (DSMB). The NCC oversees approvals of the protocol and consent forms, site compliance with regulatory issues including maintaining current IRB forms, HIPAA authorizations, and ethics certifications (human subjects training) of study personnel as required.

To provide a Comprehensive Communication System and Administrative Support: The NCC will develop and collaborate with the DMC to maintain a comprehensive communication system, arrange conference calls, draft and disseminate minutes to data analysis working group meetings, and provide reimbursement reports.

1.5.3. The Role of the National Data Management Center (DMC): The goals of the National Data Management Center (DMC) are to support the National Coordinating Center (NCC) in the scientific oversight of the trial; to facilitate the collaboration among the investigators; and to collect, manage and analyze all study data.

The specific aims of the DMC are to:

Provide Methods of Data Entry and Data Management: The DMC will implement a data entry and management system to ensure the highest quality data by developing checks for logical inconsistencies and reports to the NCC, Regional Centers, and Clinical Sites including: recruitment, retention, missing data, participant follow-up visit schedules, and treatment adherence.

Facilitate Procedures and Carry Out Quality Control: The DMC will design all data collection forms, prepare manuals of operations, train and certify all study personnel on data entry, conduct site visits, and develop systems to report serious adverse events to the appropriate organizations.

Assume Responsibility for Reporting, Statistical Design, and Analysis: The DMC will design and implement a system for random treatment assignment, develop regularly scheduled reports, including Data and Safety Monitoring Board (DSMB) reports, and conduct final and interim analyses.

1.5.4. Methods to Collect, Process and Transmit Biological Samples, Neuroimaging, and qEEG data: Standardized Collection and Quality Assurance: Each CS will follow a standardized protocol to systematically collect data used for the proposed study for clinical phenotype, neuroimaging, qEEG, and behavioral phenotype, protocol which has been used in previously published studies. Specific details of the data collection procedures for each of the markers are provided in Section 5. Imaging procedural training, and data collection will be managed by Drs. Phillips, Parsey, and Buckner. QEEG research staff training and data collection will be managed by Drs. Bruder and Pizzagalli. Each Core includes all experts in the core's biological markers and meets regularly by teleconference with Drs. Trivedi and Weissman to ensure that all biological marker-related activities are coordinated across relevant parts of the organization. Following analysis of the biological data, it will be forwarded to the DMC for data analyses for publication.

1.5.5. Approach to Biomarker Analyses: Biological markers collected in this study are divided into four main groups: 1) neuroimaging data, 2) qEEG data, 3) clinical data, and 4) behavioral data. Blood samples will also be collected from all patients, at all sites. One sample of whole blood (20ml – 3 large tubes) will be banked in yellow-top ACD (acid citrate dextrose) tubes at UTSW. In addition, blood (20 ml) will also be sent to the NIH Genetics Repository at Rutgers for tissue storage, cataloguing, and distribution for future analyses. With the advice of our consultants in genomics and proteomics, we will obtain samples of mRNA and plasma at baseline and after one and eight weeks on each treatment, for banking at the Rutgers repository, to be made available to qualified investigators under standard NIMH procedures for the repository.

Approach to Collecting Biological Markers		
	Baseline	**Week 1**
Clinical Data (n=400)	All sites, all patients	N/A
Neuroimaging* (n=400)	All sites, all patients	All sites, all patients*
Behavioral Tasks (n=400)	All sites, all patients	All sites, all patients
qEEG (n=400)	All sites, all patients	All sites, all patients
Blood collection‡ (n=400)	All sites, all patients	All sites, all patients

*DTI will not be repeated at week 1; ‡Blood also collected at week 8

2. SPECIFIC AIMS

2.1. Moderator Aims

Aim 1: To identify **baseline clinical, neuroimaging, neurophysiological, and behavioral moderators** of differential treatment outcome (mean symptom change and tolerability) for citalopram (CIT, a serotonergic antidepressant) versus placebo (PBO) for the treatment of MDD. Symptom change will be measured using mean change from baseline in the 17-item Hamilton Rating Scale for Depression ($HRSD_{17}$). Tolerability will be measured using the Frequency, Intensity, and Burden of Side Effects Rating (FIBSER) and the Systematic Assessment for Treatment Emergent Effects (SAFTEE).

Aim 1A: Clinical Moderators − To assess the extent to which baseline clinical variables – including anxious depression, melancholic depression, anger attacks, Axis II disorder, hypersomnia/fatigue, and atypical depression – will moderate treatment outcomes to CIT and PBO.

Aim 1B: Neuroimaging Moderators − To assess the extent to which the following baseline neuroimaging measures – fMRI response to two tasks (implicit emotion processing and regulation and reward processing) resting state connectivity, PASL, diffusion tensor imaging, and structural MRI measures of cortical thickness – will differentially moderate treatment outcomes to CIT and PBO.

Aim 1C: Behavioral Phenotyping Moderators – To assess the extent to which the following baseline behavioral measures – pre-treatment psychomotor slowing, cognitive control, working memory performance, and reward responsiveness – will differentially moderate treatment outcomes to CIT and PBO.

Aim 1D: Electrophysiology Moderators − To assess the extent to which the following pretreatment electrophysiology measures – resting EEG alpha and theta power, source localization measures of theta activity in the rostral ACC, and loudness dependence of auditory evoked potentials (LDAEP) will differentially moderate treatment outcomes to CIT and PBO.

2.2. Mediator Aims

Aim 2: To identify early phase (week 1) changes in **neuroimaging, neurophysiological, and behavioral tasks as mediators** of differential treatment outcomes (symptom change, tolerability) to CIT and PBO.

Aim 1A: Neuroimaging Mediators − To assess the extent to which changes in the following neuroimaging measures – fMRI response to two tasks (implicit emotion processing and regulation and reward processing), resting state connectivity, and PASL measures of regional cerebral blood flow from baseline to one week posttreatment – will differentially mediate treatment outcomes to CIT and PBO.

Aim 1C: Behavioral Phenotyping Mediators – To assess the extent to which changes in psychomotor speed, cognitive control, working memory performance, and reward responsiveness from baseline to one week posttreatment – will differentially mediate treatment outcomes to CIT and PBO.

Aim 1D: Electrophysiology Mediators − To assess the extent to which changes in the following electrophysiology measures – resting EEG alpha and theta power, theta activity in the rostral ACC, and loudness dependence of auditory evoked potentials (LDAEP) from baseline to one week posttreatment – will differentially moderate treatment outcomes to CIT and PBO.

2.3. Main Treatment Effects Aim

Aim 3: To compare the 8-week outcomes of CIT vs. PBO based on rates of **symptom remission**, defined as a score of 7 or less on the 17-item Hamilton Rating Scale for Depression ($HRSD_{17}$).[54]

2.4. Methodological Innovation Aims

Aim 4: To develop methods to examine the complex interactions of clinical, neuroimaging, neurophysiology, and behavioral moderators and mediators of differential treatment outcomes (mean symptom change, tolerability) for CIT and PBO. We provide an initial example of four possible complex interaction aims as illustrations. Secondary analyses will be conducted to explore further interactions.

Aim 4A: To assess the interaction between fatigue and psychomotor retardation and baseline imaging measures of activity and connectivity in reward processing neural systems, and task-specific reaction times, to differentially moderate CIT vs. PBO treatment outcomes.

Aim 5: To develop a **depression differential treatment response index (DTRI)** that integrates moderators across clinical and biological variables. In developing markers for inclusion in the DTRI, we propose that any individual marker must independently or interactively predict treatment response. **Based on available literature, a model that can provide 20% improvement using DTRI selected treatments over the currently used nonspecific treatment selection approach could signify a clinically meaningful treatment prediction**.[55]

Aim 6: To develop a composite scale of **treatment acceptability** (which incorporates depressive symptom outcomes, treatment tolerability, and the patient's functional status) and compare 8-week outcomes for CIT and PBO.

Aim 7: To **explore the clinical, neuroimaging, electrophysiology, and behavioral phenotype biosignatures of differential treatment outcome** identified for first step responders, among second step treatment responders (i.e., first step nonresponders).

3. BACKGROUND

3.1 Burden of MDD: Depressive disorders are one of the leading causes of disability-adjusted life years, worldwide.[56] According to the World Health Organization (WHO) depressive disorders are the fourth leading cause of disability-adjusted life years, worldwide and by 2020 are estimated to be second only to ischemic heart disease.[57] This debilitating illness will affect up to 16.2% of Americans at some point in their lifetime.[58]

3.1.a Low Remission Rates: The goal of treatment for those suffering from MDD is remission (i.e., the absence of depressive symptoms),.[59] and yet, two-thirds of patients treated with a first step antidepressant do not achieve remission of symptoms[4] and successive treatment steps lead to diminishing remission rates.[23] Moreover, large numbers of patients either discontinue treatment prematurely due to side effects, or become discouraged and drop out of treatment altogether.[4] Advances in antidepressant drug treatment discovery have been very limited, and simply waiting for the next big breakthrough is certainly not a sure bet.[60] In fact, disease heterogeneity[61] is likely to complicate pharmacological approaches going forward, even in face of novel drug development. Establishing methods for identifying which patient is likely to respond (and have fewer side effects) to current treatment options is an essential priority for our field.[62]

3.2 Personalized Treatment: Studies of cardiovascular disease, asthma, breast cancer, lung cancer, multiple sclerosis, macular degeneration and other medical illnesses have been successful in identifying important moderators of treatment response, leading to the development of personalized treatment approaches.[24-29] Given the ineffectiveness of first-step treatments for MDD, there is a clear and urgent need to identify factors that can be used to individualize treatment (i.e., markers that maximize effectiveness and minimize risk for toxicity).[53, 62] Personalizing treatments for MDD is likely accomplished by utilizing an array of clinical and biological markers that differentiate treatment selection for individual patients.

3.3 Rationale for Selection of Markers for EMBARC: Most previous research has evaluated clinical and biological markers for treatment response for individual treatments. While this research provides indicators of those likely to an active treatment, the context of these studies does not allow us to determine which, if any, of these markers is specific to active treatment, and which are nonspecific markers of placebo response. The clinical and biological markers that we have selected for inclusion in this study are based on a thorough review of the relevant literature. The major advantage of our current approach stems from the ability to account for placebo response as well as evaluate a myriad of clinical and biological markers and their interactions in a single study. In addition, this study will allow us to establish standard procedures for the collection, processing, data management, and analysis of a combination of biological and clinical markers. As such, this study will set a precedent for the field for future definitive biosignature endeavors.

3.4 Clinical Moderators:

3.4.a. Demographic Moderators (gender and employment status): An analysis of gender differences in antidepressant response with either sertraline or imipramine[63] showed that men were more likely to respond to the nonSSRI, imipramine, while women were more likely to respond to the SSRI, sertraline. Similar comparisons between SSRIs and BUP have been conducted from our group, with less definitive results with regard to gender, although suggesting an advantage of SSRIs over BUP in anxious women.[64]

3.4.b. Early Life Trauma and Life Events: Trauma history. Early life trauma is a well established risk factor for the development of depression.[65-67] Whether such trauma affects the likelihood of response to antidepressant treatment has not been adequately studied. In a post-hoc analysis of a clinical trial comparing nefazodone with a specific form of cognitive behavioral therapy, though neither treatment showed overall superiority, those subjects with a history of childhood trauma showed a superiority of psychotherapy alone to nefazodone, suggesting that a history of childhood trauma may be a moderator for antidepressant response.[68] Since most studies of antidepressant medication treatment have not carefully collected histories of childhood trauma, this may present an important lead, which should be pursued in a study of putative moderators of antidepressant treatment outcome. Furthermore, number of negative life events also predicted differential response with patients, with greater numbers of negative life events doing better in cognitive-behavioral therapy (CBT) than with antidepressants.[69]

3.4.c. Symptom Features (melancholic depression, anxious depression, atypical depression, depression with anger attacks, depression with hypersomnia/fatigue): Approximately 1/3 of MDD outpatients present with melancholic features.[70] Level 2 STAR*D findings from our group[34] showed that the odds ratio of remission were 0.4 for BUP vs. 0.6 for sertraline in melancholic depression (vs. nonmelancholic depression), supporting the view, based on a previous study from our group[71] that SSRIs are mildly less effective in this depressive subtype.

Analyses of STAR*D data suggest that atypical MDD is predictive of poorer outcome to treatment with an SSRI antidepressant.[72] However, with a multivariate analytic approach, atypical depression was no longer predictive once chronicity was accounted for.

In Levels 1 and 2 of the STAR*D study,[73] in a sample of 2,876 adult outpatients with MDD, patients with high levels of anxiety (anxious depression) were significantly less likely and took longer to remit than those with nonanxious depression. Ratings of side effect frequency, intensity, and burden, as well as the number of serious adverse events, were also significantly greater in the anxious depression group, as was the rate of hospitalization for nonpsychiatric reasons. Similarly, in Level 2, patients with anxious depression fared significantly worse in both the switching and augmentation options. Based on a metaanalysis of 10 pooled double-blind studies, our group has found a modest benefit for SSRIs over BUP (6% benefit in response rates) in patients with anxious depression.[10]

Another common subtype (approximately 30 to 40% prevalence in depressed patients) is depression with anger attacks, which are sudden intense spells of anger.[36] In studies from our group, depressed patients with anger attacks showed a significantly blunted prolactin response to fenfluramine challenge compared to depressed patients without anger attacks, suggesting that depressed patients with anger attacks may have a relatively greater serotonergic dysregulation than depressed patients without these attacks.[74] Our studies also showed that anger attacks have disappeared in 53% to 71% of depressed patients treated with serotonergic antidepressants.[36] This is consistent with the observation that SSRIs show anti-aggressive properties in several animal models.[37] On the other hand, BUP appears to actually enhance aggressiveness in mice[75] and paradoxical aggression may be precipitated by the use of noradrenergic drugs in MDD.[76] A metaanalysis of CBT studies[77] suggests that CBT has a smaller effect size in depression with anger attacks than in depressed patients without these attacks.

Sleep disturbances are commonly reported in patients with MDD, with some patients reporting insomnia, while others report a picture of fatigue and hypersomnia.[78] For patients within the latter group, BUP was found to be superior to SSRI in a pooled analysis.[79]

3.4.d. Chronic Depression: While some studies have not supported chronicity as a moderator of outcome,[80, 81] analyses of the STAR*D dataset strongly support chronicity as predictive of treatment outcome,[82, 83] as do previous analyses, which suggest that this may be caused by a lower nonspecific or placebo response in the most chronic patients.[84-86]

3.4.e. Family History: A family history of severe psychiatric illness (depression requiring hospitalization, psychosis and others) has long been thought to be a predictor of response in depression.[87] However, recent data from STAR-D suggest that once other variables are accounted for, a family history of depression (severity not specified) is not a robust predictor.[88]

3.4.d. Comorbidity Moderators: Depressive illnesses are often comorbid with personality disorders. SSRIs have shown to be effective in treating affect lability, impulsivity, and aggression in patients with borderline personality disorder in at least 10 open studies and two double-blind studies of SSRIs.[89] A recent placebo-controlled trial from our group found that patients with MDD superimposed on Axis II disorders did better on the SSRI paroxetine than they did in CBT, whereas MDD patients free from Axis II disorders showed the opposite pattern.[69] No definitive studies with Axis II comorbidities and BUP have been conducted, however, given the observation of paradoxical worsening with the use of noradrenergic drugs in MDD with borderline features,[76] one may expect MDD patients with axis II disorders to do less well with BUP.

3.4.e. Other Possible Moderators. The literature on clinical predictors of treatment outcome has been recently reviewed.[67] In addition to the variables listed above there are other candidate variables with fewer data or with conflicting studies. These include medical comorbidity[90]; hopelessness levels; physical symptoms of depression; subsyndromal bipolar symptoms; and psychomotor retardation. Social support, and the quality of marital and family relationships have also been linked to outcome in MDD.[91] The greatest support for demographic and clinical variables other than those listed comes from Level 1 of the STAR*D study.[82] In that sample, participants who were Caucasian, female, employed, or had higher levels of education or income had higher HAM-D remission rates. Conversely, longer index episodes, more concurrent psychiatric disorders (especially anxiety disorders or drug abuse), more general medical disorders, and lower baseline function and quality of life were associated with lower HAM-D remission rates. While these variables have had some support in other studies, none has been clearly replicated or firmly established as a predictive variable. Moreover, the STAR*D investigators have not conducted multivariate analyses of their data set to determine which of these inter-correlated variables might contribute unique variance to the estimation of treatment outcome. In a prospective trial, each of these is worth assessing as they are easily determined and may interact in as yet unknown ways with biologic variables, including genetic constitution and traumatic life events.

3.4.e. Summary of Clinical Moderators: Most predictive research to date has focused on assessing whether clinical and demographic factors are associated with treatment outcomes for a single antidepressant. Unfortunately, to date this approach has been met with limited success,[32] and has led to few practical results, and to the best of our knowledge, none that have been used to guide personalized care for the treatment of depression. In summary, we aim to find specific clinical markers that are likely to differentiate active treatment with CIT from placebo for Stage 1 study participants.

3.5. Neuroimaging Moderators/mediators:

Core MDD symptom dimensions include anxiety, low mood and anhedonia. Neuroimaging studies in MDD have linked these symptoms to dysfunction in several distributed neural circuits, which include subcortical systems involved in emotion and reward processing (e.g. amygdala, ventral striatum), and medial prefrontal regions involved in both the processing and implicit regulation of emotion. While no studies have yet examined the extent to which neuroimaging measures may act as moderators of *differential* response to different treatments in MDD, numerous prior studies have examined neural predictors of outcome to a range of treatments. Predictor studies have implicated a role for each these neural systems, but in particular focused on the role of the anterior cingulate cortex (ACC) and medial prefrontal cortex (mPFC). It has been suggested that there may be a consistent relationship between pretreatment ACC/mPFC and treatment response, which may differ in sign depending on whether a purportedly "bottom-up" medication approach or "top-down" psychotherapy approach is taken.[23] The neuroimaging paradigms selected for the proposed study will examine a range of relevant neural systems which involve the ACC/mPFC, thus maximizing both the scientific impact of the proposed study and our ability to find robust neuroimaging biosignatures of treatment for MDD. Furthermore, our neuroimaging paradigms are focused to examine neural circuitry supporting key information processing domains highlighted in the new **Research Domain Criteria (RDoC) for mental disorders**, including **fear/emotion and reward functioning**,[92] and pave the way for personalized treatment approaches based on an enhanced understanding of underlying pathophysiology in MDD. Critically, we are applying a systems-level exploration that includes structural MRI analysis, task-based analysis that targets known neural circuits using well develop fMRI paradigms, and analysis of intrinsic functional (effective) connectivity, together with measures of white matter integrity, in these neural circuits, that have more recently emerged as powerful tools for assessing the integrity of brain systems and differences among individuals. We took into consideration the putative mechanisms of action of the interventions proposed in selecting our neuroimaging approaches.

3.5.a. Neuroimaging Moderator Assumptions: Each of these assumptions examines different, but partially overlapping and complementary, neural systems, measured pretreatment.

Assumption 1: Medial prefrontal-limbic network. Activation of the ACC/mPFC and its functional/effective connectivity with the amygdala during the **emotional conflict task**, a probe of both emotional processing and implicit (i.e. nonconscious) emotion regulation, will differentiate between response to CIT versus PBO in the first eight week period of the trial (Stage 1), and to a lesser extent, will help distinguish between CIT versus BUP in the second eight week period of the trial in MDD nonresponders to either PBO or CIT from the first eight-week period of the trial who are re-assigned to CIT or BUP (Stage 2). Medial prefrontal-subcortical connectivity during **resting-state-fMRI** will also differentiate between response to CIT versus PBO, and to a lesser extent, will help distinguish between CIT versus BUP in MDD nonresponders to either PBO or CIT from the first eight-week period of the trial who are re-assigned to CIT or BUP (Stage 2). The integrity of white matter, as measured by diffusion tensor imaging, in medial prefrontal-subcortical circuitry will be an additional measure that, together with the above measures of activity, functional/effective connectivity and resting state connectivity in medial prefrontal-subcortical circuitry, will help differentiate between response to CIT versus PBO, and, to a lesser extent, between response to CIT versus BUP in Stage 2.

Assumption 2: Reward network. The following reward circuitry neuroimaging measures during the **reward processing task** will help differentiate between response to BUP vs. CIT among Stage 2 participants in the second eight week trial period (although with smaller effect sizes than neuroimaging measures in Stage 1 because of the pre-exposure of Stage 2 participants in Stage 1 to either CIT or PBO). These measures include: activity in, and functional/effective connectivity between, ventral striatal and orbitomedial prefrontal cortical regions; connectivity between the ventral striatum and orbitomedial prefrontal cortex during **resting-state-fMRI;** integrity of white matter, as measured by diffusion tensor imaging, in ventral striatal-orbitomedial prefrontal cortical circuitry.

We outline below additional evidence supporting these assumptions, divided by neural system.

3.5.a.1. Medial prefrontal-limbic network: It is now well-established that emotion processing and implicit (nonconscious) emotion regulation neural systems, centered on amygdala and orbitomedial prefrontal cortical (OMPFC) regions are modulated by serotonin (5HT) neurotransmission in a wide variety of tasks.[93-98] Likewise, SSRI medications modulate emotion-induced activity in this network in MDD,[99, 100] and response to SSRIs are predicted by activity there.[100-103] Less is known about the effects of nonserotonergic medications on this network. We will examine the function of this network during emotional processing and implicit regulation with the **emotional conflict task** in order to differentiate between response to CIT and PBO. The function of the medial prefrontal cortex can also be examined by examining low-frequency (< ~0.1 Hz) BOLD fluctuations (LFBF) between interconnected brain regions using functional connectivity during resting-state fMRI (fc−fMRI; see[205]). The measures therefore provide a complementary view of this neural circuitry as the task-based activation measures. Depressed MDD individuals show resting state LFBFs that are abnormally increased in the self-referential default network,[104] but abnormally reduced between subcortical regions and ACC,[105, 106] with the subcortical-ACC LFBF increasing after the SSRI sertraline.[105] Convergent evidence for a medial prefrontal default mode abnormality in MDD comes also from analysis of task-independent deactivations.[107] We will therefore examine connectivity of this network using **resting-state fMRI**, in order to provide complementary data on measures of medial prefrontal-subcortical function pre-treatment that may moderate differential response to CIT versus PBO.

3.5.a.2. Reward Network: Animal and human neuroimaging studies have shown that reward processing neural circuitry, centered on ventral striatum and interconnected OMPFC, are modulated by dopamine. Dopamine modulates neuronal activity in the ventral striatum in rodents during the learning of reward-related behaviors.[108-111] Pharmacological fMRI studies in humans indicate that increasing dopamine "equalizes" the magnitude of ventral striatal activity during anticipation of both reward and punishment.[112] Likewise, patients with MDD have abnormally blunted ventral striatal responses to reward stimuli.[113] Consistent with these findings, we recently found that acute administration of the dopaminergic agent BUP during a reward task led to normalization of ventral striatal hypoactivation in MDD.[41] We will include **a reward-processing task** that will measure activity and connectivity in dopamine-modulated reward processing ventral striatal-OMPFC neural circuitry during anticipation, and also receipt, of possible win or loss, and hypothesize that it will differentiate between response to dopaminergic BUP and serotonergic CIT in Stage 2, although with smaller effect sizes than neuroimaging measures differentiating between response in Stage 1, because of the pre-exposure in Stage 1 of Stage 2 participants to either CIT or PBO.

3.5.a.3. We will also examine connectivity during **resting-state fMRI (low frequency BOLD fluctuation and pulsed arterial spin labeling (PASL))** between these regions to provide complementary data on measures of ventral striatal-OMPFC function pre-treatment that may moderate differential response to CIT versus PBO, and, to a lesser extent, differential response to CIT versus BUP in Stage 2.

3.5.a.4. Diffusion tension imaging (DTI): DTI is a MRI based method that can measure the macroscopic organization of axons in the living human brain.[114] We will use DTI to provide complementary data on the integrity of WM tracts in reward circuitry pre-treatment that may moderate differential response to the two different medications. It measures the organization, or lack of free diffusion, of water molecules in tissue. For example, the water molecules in an axon cannot move as freely as those in cerebrospinal fluid (CSF). One of the most common outcome measures used in DTI analysis is fractional anisotropy (FA), a measure of the degree of disorganization of the water molecules and, by inference, the fibers in the brain. Lower FA can be due to changes in the density of the axons, axonal diameter, myelination, coherence of the fiber tract, or localized water content. Generally, we consider lower FA to reflect a larger disorganization of fibers. In a recent meta-analysis of DTI data in mood disorders, 21 of 27 studies found significantly lower FA in mood disorders subjects compared to healthy volunteers.[115] To our knowledge only one published study has used FA to predict treatment response, which was conducted in late-life depression. The study showed that response to sertraline was associated with lower frontal FA values.[116]

Probabilistic tractography is a second DTI data analytic technique that we will use to examine the strength of WM connectivity between a given seed region of interest (e.g., amygdala) and other regions across the whole brain. These DTI measures will provide complementary data on the integrity of WM tracts in medial prefrontal-subcortical circuitry pre-treatment that may moderate differential response to the two different treatments.

3.5.a.5. Using Cortical Thickness to Predict Treatment Remission. We have also recently explored the possibility that cortical thickness might predict response to antidepressant treatment, as cortical thickness is an index of neuronal integrity and arborization.[117, 118] Our colleagues have found lower neuronal density

postmortem in the ventral and dorsal lateral prefrontal cortex in MDD.[119] High resolution T1 structural MRI scans were obtained on a GE 3T magnet in 20 MDD subjects scanned during a MDE. Cortical thickness was measured for 10 remitters and 10 nonremitters all treated with ESC. The measurement of the thickness is performed using Freesurfer software (see below).[117] Several left frontal cortical regions were smaller in nonremitters compared to remitters. The smallest significant difference was in the left inferior opercular region (5% smaller, p=0.017). The largest effect was in the left fronto-marginal cortex (12% smaller, p = 0.009). We hypothesize that nonresponders to CIT will have thinner left frontal cortical thickness than responders to CIT.

3.5.b. Neuroimaging Mediator Assumptions: Each of these assumptions examines changes in the above measures of activity, functional/effective connectivity, resting state connectivity and white matter integrity in the above two key neural networks measured between pre-treatment and one week after initiation of treatment.

Assumption 1: Medial prefrontal-limbic network. Changes from pre-treatment to one week after treatment initiation in the above fMRI and resting state connectivity measures during performance of the **emotional conflict task**, will differentially mediate response to CIT versus PBO, and, to a lesser extent, response to CIT compared to BUP in Stage 2 for MDD nonresponders.

Assumption 2: Reward network. Changes from pre-treatment to one week after treatment initiation in the above fMRI and resting state connectivity measures during performance of the **reward task**, will differentially mediate response to CIT in Stage 1 compared to BUP in Stage 2 for MDD nonresponders to CIT. These changes in reward network neuroimaging measures from pre-treatment to one week post treatment initiation will, however, differentially mediate response to CIT versus BUP in Stage 2 participants with smaller effect sizes than those neuroimaging measures differentiating between response in Stage 1 because of the pre-exposure in Stage 1 of subsequent Stage 2 participants to either CIT or PBO.

3.6. Electrophysiology Moderators/mediators:

Quantitative EEG and Response to Antidepressants. EEG measures of brain activity in a resting state may provide a cost effective predictor of antidepressant response. MDD patients who responded to imipramine had significantly less pretreatment theta power compared to nonresponders.[44] Pretreatment theta power at frontal or temporal sites (referenced to Fpz) was also lower in responders to SSRI or venlafaxine than in nonresponders.[43] These scalp findings contrast with recent data highlighting a robust link between *increased* pre-treatment resting theta activity in the rostral anterior cingulate cortex (rACC) and better response to nortriptyline,[49] citalopram,[48] reboxetine,[48] fluoxetine, or venlafaxine,[47] which are summarized below. A recent report suggests that an ATR index could predict response to escitalopram or bupropion[120] but this needs independent replication. Others found increased posterior alpha power in patients who responded to amitriptyline[121] or fluoxetine.[120] In a recent replication[42], greater alpha in treatment responders than nonresponders was present both for patients treated with an SSRI or dual therapy (SSRI plus bupropion, duloxetine, or venlafaxine). This difference in alpha had a large effect size (d= 1.17) and was able to predict treatment response with high positive predictive value (89%) and specificity (85%), but lower sensitivity (57%). Based on these findings, we hypothesize that increased pre-treatment alpha will predict positive response to CIT but not PBO.

Resting rACC activity and Response to Antidepressants. In 1997 Mayberg et al.[101] first reported that increased resting metabolism in the rACC predicted better antidepressant response in MDD. We recently performed a meta-analysis to assess the robustness of this association and found that, across 23 studies involving 426 MDD subjects, the mean weighted effect size for a link between increased pre-treatment rACC activity and better treatment response was 0.92 (95% CI: 0.44-1.39; Z=3.78, p<.001), reflecting a large effect size (Pizzagalli, D.A., *in press*. Frontocingulate dysfunction in depression: Towards biomarkers of treatment response. Neuropsychopharmacology Review). Measuring resting EEG data before an open-label trial with nortriptyline, Pizzagalli et al.[49] first replicated Mayberg's findings using EEG source localization. Specifically, increased resting pre-treatment rACC activity predicted better response to nortriptyline 4-6 months later. Recent receiver operating characteristic analyses showed that resting rACC theta activity correctly classified 89% of eventual responders and 89% of nonresponders, highlighting a strong sensitivity/specificity profile (Pizzagalli, in press). These data were recently replicated in two studies using the same source localization technique. Mulert et al.[48] reported that MDD subjects responding to either SSRI (citalopram) or norepinephrine reuptake inhibitor (reboxetine) treatment had higher pre-treatment theta rACC activity than nonresponders. The effect size differentiating groups (d=1.33) was similar to ours (d=1.43). In a placebo-controlled study, Korb et al.[47] found that drug responders had significantly higher rACC theta activity than nonresponders. Of note, findings (1) were specific to drug (vs. placebo) responders, (2) emerged when considering both an SSRI

(fluoxetine) and a SNRI (venlafaxine), and (3) were based on the rACC region emerging from Pizzagalli et al.[49] These data suggest that the link between theta rACC activity and treatment response is specific to antidepressant and not placebo responses. Moreover, this finding has emerged for a variety of SSRIs but not bupropion,[122] suggesting that the rACC marker might be used to differentially predict response to citalopram vs. bupropion Based on these findings, we hypothesize that increased pre-treatment rACC theta activity will predict positive response to CIT but not PBO (Stage 1) or BUP (Stage 2).

Loudness Dependency of Auditory Evoked Potentials as an Index of Central Serotonergic Activity. Auditory evoked potentials (N1-P2) may provide a marker of central serotonin activity, which could identify patients with a serotonergic deficit responsive to SSRI.[123] Primary auditory cortex has extensive serotonergic innervation[124, 125] and N1-P2 potentials are known to be generated in this region.[126] Findings suggest that loudness dependence of auditory evoked potentials (LDAEP), the increase in N1-P2 amplitude with increasing tone intensity, provides a noninvasive indicator of central serotonergic activity. Direct evidence of an inverse relationship between serotonergic activity and LDAEP has been documented in recordings from primary auditory cortex in cats.[127] Also, patients with low serotonergic activity, as evidenced by pronounced loudness dependence before treatment, responded better to SSRI than patients with evidence of high serotonergic activity.[50-52] In contrast, reduced loudness dependence was associated with greater improvement in depressive symptoms following treatment with the selective noradrenaline reuptake inhibitor reboxetine.[128] Moreover, response to citalopram was characterized by strong LDAEP, whereas response to reboxetine was characterized by weak LDAEP.[50] Based on these findings, we hypothesize that potentiated pre-treatment loudness dependence will predict better response to CIT but not PBO.

3.6.a. qEEG Moderator Assumptions

Assumption 1 (moderator): Increased pre-treatment alpha over posterior regions will predict positive response to CIT but not PBO in the first eight week period of the trial (Stage 1).

Assumption 2 (moderator): Increased pre-treatment rACC theta activity will predict positive response to CIT but not PBO in the first eight week period of the trial (Stage 1), and to citalopram but not BUP in the second 8-week period of the trial (Stage 2).

3.6.b. LDAEP Moderator Assumptions

Assumption 1 (moderator): Potentiated pre-treatment loudness dependence will predict better response to CIT but not PBO.

3.6.c. qEEG Mediator Assumptions

Assumption 1 (mediator): No change in the above resting EEG measures from pretreatment to one week after initiation of treatment will predict poor response to CIT.

3.6.d. LDAEP Mediator Assumptions

Assumption 1: No change in the above LDAEP EEG measure from pretreatment to one week after initiation of treatment will predict poor response to CIT.

3.7. Behavioral Phenotypes Moderators/mediators:

We propose to use behavioral phenotyping/neurocognitive tasks that: (1) have shown most promise of discriminating antidepressant responders and nonresponders and identifying moderators of differential response to CIT vs. BUP; (2) probe important domains identified within NIMH's new Research Domain Criteria (RDoC) for mental disorders (http://www.nimh.nih.gov/research-funding/rdoc.shtml; Insel and Cuthbert[92]); and (3) maximize integration with the proposed fMRI component. In addition, task selection was guided with the aim of minimizing patients' burden. Based on these considerations, we propose to investigate Psychomotor Slowing, Executive Function (Cognitive Control), Working Memory, and Reward Responsiveness across five different tasks (Table 1).

Table 1

Domain/Construct	Task	RDoc	Hypotheses CIT (vs. placebo)
Psychomotor Slowing	Choice RT task Word Fluency	No	Psychomotor slowing will predict CIT *non*response
Cognitive Control	Flanker task	Yes	Reduced post-error adjustments (Laming effect) will predict CIT *non*response
Working Memory	A Not B task	Yes	Reduced working memory will predict CIT *non*response.
Reward Responsiveness	Probabilistic Reward task	Yes	Reduced reward responsiveness will predict CIT *non*response

3.7.a. Psychomotor slowing: Several studies have shown that psychomotor slowing is a specific predictor of SSRI treatment nonresponse[85, 129, 130] (see also Caligiuri et al.[131]). For example, using a brief word fluency task (the Controlled Oral Word Association Test), Taylor and colleagues[85] reported that psychomotor slowing was the strongest predictor of fluoxetine nonresponse 12 weeks later (see also Flament et al.[129]). Because psychomotor slowing is an important component of melancholia, which has been linked to dopaminergic dysfunctions, it has been hypothesized that patients with psychomotor slowing might respond better to treatments targeting dopamine neurotransmission.[85] In line with this assumption, patients with retarded depression treated with antidepressants over 6 weeks showed graded improvements based on the degree of dopamine affinity[132]: patients treated with amineptine (a selective inhibitor of dopamine reuptake) showed greater improvement on both motor retardation and depression severity compared to patients treated with minaprine, clomipramine, or placebo. Interestingly, low pre-treatment psychomotor speed predicted better response to an 8-week treatment with bupropion sustained release (150 mg/day) in MDD outpatients.[133] Based on these findings, we hypothesize that pre-treatment psychomotor slowing will predict poor response to citalopram, but not placebo (Table 1).

3.7.b. Cognitive Control (post-error behavioral adjustments). As summarized in the EEG section, decreased pre-treatment activity within the rACC has been reliably found to predict poor antidepressant response. Of relevance to the current proposal, pre-treatment rACC function predicted response to SSRIs (paroxetine: Kennedy et al.[134], Saxena et al.[135]; citalopram: Mulert et al.[48]; escitalopam: Langenecker et al.[136]; fluoxetine: Korb et al.[47]) but not bupropion (Little et al.[122]) or placebo (Korb et al.[47]). Studies from Dr. Pizzagalli's laboratory have shown that abnormal reactions to errors in dysphoric subjects,[137] unmedicated subjects with MDD,[138] and psychiatrically healthy individual carrying genetic variants associated with increased depression risk[139] are associated with dysfunctional activity in the rACC. Importantly, among controls, resting rACC activity predicted post-error behavioral adjustments during a Flanker task.[137] Based on these findings, we propose that post-error behavioral adjustments in executive tasks (e.g., Flanker task) can be used as a behavioral probe of rACC function. We hypothesize that reduced post-error behavioral adjustments will predict poor outcome to CIT, but not PBO. Recent findings that Stroop performance predicted treatment response to fluoxetine in patients having a major depressive disorder[85] and to sertraline but not CBT in patients with multiple sclerosis and depression are consistent with this hypothesis.[140]

3.7.c. Working memory The Columbia group found global deficits on neuropsychological tests were predictive of poor response to an SSRI, with deficits being greatest in the working memory domain.[141] Moreover, performance on working memory tasks that require manipulation of information was found to improve following SSRI treatment.[142] We therefore predict that performance on the A NOT B working memory test[141] will be predictive of response to citalopram.

3.7.d. Reward Responsiveness. Anhedonia has been found to predict (1) depressive symptoms (e.g., Hundt et al.[143]); (2) time to recovery (e.g., McFarland et al.[144]); (3) poor outcome (e.g., Kasch et al.[145]; Spijker et al.[146]); and (4) chronic course of depression.[147] In the current research, we propose to investigate a core behavioral component of anhedonia – reward responsiveness – which will be objectively assessed using a probabilistic reward task. Critically, performance in this task has been found to (1) correlate with current and predict future anhedonic symptoms (e.g., Pizzagalli et al.[148, 149]), (2) be modulated by dopaminergic compounds (e.g., Pizzagalli et al.[150]), and (3) correlate with brain activation in reward-related striatal regions (e.g., Santesso et al.[151]). Based on these findings, we hypothesize reduced reward responsiveness will predict poor CIT response in Stage 1 but better response to BUP in Stage 2.

3.7.e. Behavioral Moderator Assumptions:

Assumption 1 (Stage 1): Pre-treatment psychomotor slowing will predict poor response to CIT, but not PBO.

Assumption 2 (Stage 2): Pre-treatment psychomotor slowing will predict positive response to BUP.

Assumption 3 (Stage 1): Reduced post-error behavioral adjustments will predict poor outcome to CIT, but not PBO.

Assumption 4 (Stage 2): Reduced post-error behavioral adjustments will predict poor outcome to CIT.

Assumption 5 (Stage 1): Better performance on A, not B working memory test will predict positive CIT response, but not placebo.

Assumption 6 (Stage 1): Reduced reward responsiveness will predict poor outcome to CIT, but not placebo.

Assumption 7 (Stage 2): Reduced reward responsiveness will predict positive response to BUP.

. **3.7.f. Behavioral Mediator Assumptions:**

Assumption 1: No change (no increase) in post-error adjustments, psychomotor slowing, working memory, or reward responsiveness from pre-treatment to one week after treatment initiation will predict poor outcome to CIT.

3.8. Depression Differential Treatment Response Index (DTRI): If an array of significant biomarkers for depression could guide response and tolerability by dose for select antidepressant medications, we could radically modify clinical practice. The enduring aim of this RFA is designed to develop a framework to test whether differences in various biomarkers moderate and mediate differences in treatment outcome. Our ultimate goal is to develop a depression differential treatment response index (DTRI), similar to the Framingham Risk score for cardiac risk that personalizes treatment selection for individual depressed patients by incorporating significant predictors of differential treatment response which maximize symptom reduction and treatment tolerability. For example, a recent study has demonstrated that microarray gene expression profiles from a panel of 10 different human osteosarcoma xenografts may predict their sensitivity to ifosfamide, doxorubicin, and cisplatin.[152]

There are two notable studies that have developed composite indices to predict treatment outcome with MDD. The first study, the Munich Antidepressant Response Signature study (MARS),[153] enrolled 842 psychiatric inpatients with MDD and bipolar disorder to identify the extent to which demographic factors, clinical history (disease characteristics, current episode severity, comorbidity, somatic comorbidity) and degree of HPA-axis dysregulation (CRH) predict response and nonresponse to a variety of antidepressant medications.[153] Though a composite index was developed for the prediction of nonresponders, the study was observational and naturalistic permitting the choice and length of treatment up to the clinician. As such, subjects took a variety of medications and were discharged before attaining remission. The second study, the Genome-based Therapeutic Drugs for Depression study (GENDEP),[154] has examined the influence of gene variation for encoding key proteins in serotonin, norepinephrine, neurotrophic and glucocorticoid signaling to determine treatment response to a SSRI and a SNRI. The GENDEP study enrolled 760 adult patients with MDD who were treated with either escitalopram or nortriptyline in a 12 week open-label part-randomized multicenter study, examining 10 candidate genes, including 5-HTLPPR, HTR2A, NET, TREK1, and 2D6. Results indicated that HTR2A significantly predicted response to escitalopram, and SLC6A2 predicted response to nortriptyline, and NR3C1 predicted response to both antidepressant medications.[154] The main finding was that single marker associations explained only a small proportion of the variance to these antidepressant medications, supporting the logic for a multivariate approach to predicting treatment response. We know of no studies that combine biomarkers to evaluate likelihood of higher treatment response and or tolerability to antidepressant medications. As such, our focus is on simultaneous analyses of both efficacy and tolerability to develop the composite index.

3.9. Goals of the Current Study: This study embarks on a novel initiative, which compares an active antidepressant treatment with placebo, as well as a second step switch for nonresponders to either treatment, and for the first time assesses clinical, neuroimaging, neurophysiological, and behavioral moderators (and the interrelationship of each group) of treatment outcome. It is our intention to develop a predictive model that integrates the best current evidence to provide personalized treatment, which can be further modified as new evidence emerges.

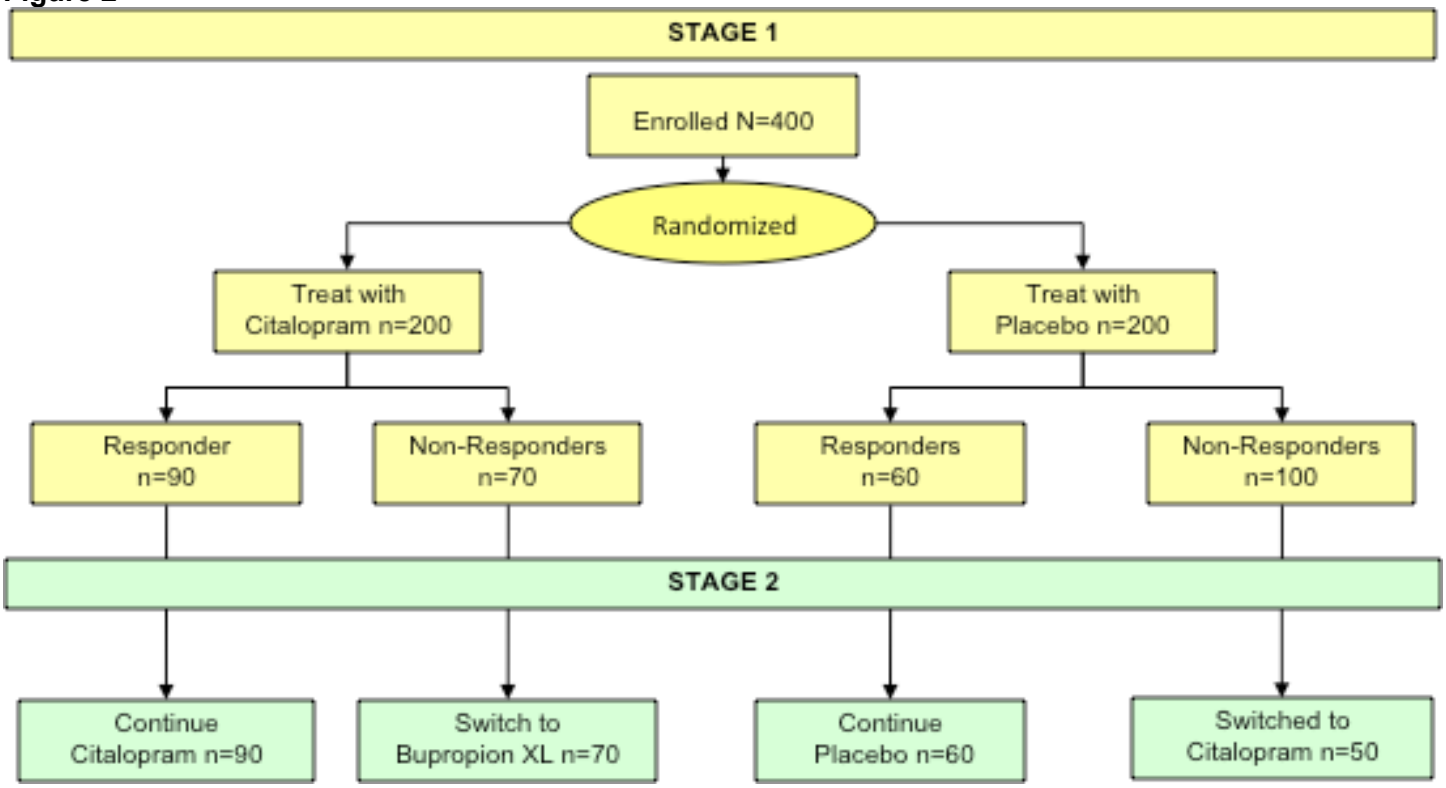
3.10. Significance and Importance: This study represents, to the best of our knowledge, the first attempt to construct a comprehensive biosignature for the personalized care of MDD. While necessary to choose two or more distinct treatments to construct the personalized care model, the study has greater implications for future studies because: 1) it details, for the first time, *how* to collect, process, and analyze a wide array of clinical and biological markers in the context of a clinical trial, which can be extended to other competing treatments and disorders; 2) it has the potential to produce, for the first time, a clinically useful metric to guide treatment; 3) uses a systematic process to choose the best clinical, neurobiological and neurophysiological moderator candidates, a process that can be extended to other psychiatric disorders; and 4) the extensive clinical and biomarker database can be used as a national resource for other investigators in the future as the field develops to identify other potential moderators.

4. PRELIMINARY STUDIES

Preliminary studies supporting the choice of specific markers and paradigms included in the study are provided in supporting Appendix I.

5. RESEARCH DESIGN
5.1 Study Overview

Figure 2



5.1.a. Design: The current study is designed to identify biomarkers for the prediction of differential treatment outcomes between the SSRI antidepressant citalopram (CIT) and placebo (PBO) in a randomized trial for patients with MDD. In addition, a second stage will collect data to explore moderators and mediators of treatment outcomes between pharmacologically distinct active treatment arms: citalopram (CIT), a serotonergic antidepressant or bupropion (BUP), a nonserotonergic antidepressant. To reduce biologic heterogeneity, we will only enroll patients with early onset of DSM IV MDD (before age 30) because these criteria in probands have been shown to be associated with increased familial loading in families. Patients will also have either chronic MDD (episode duration ≥2 years) or recurrent MDD with 2 or more recurrences (including current episode). Additionally, patients will be required to have a current symptom severity score of 20 on the Montgomery Asberg Depression Rating Scale.[155] In the first stage, patients will receive an 8−week course of treatment among in of the two study arms. As part of the SMART design patients that have not achieved response at the end of 8 weeks to their stage one treatment, defined by < 50% improvement on the 17-item Hamilton Depression Rating Scale ($HRSD_{17}$)[156, 157], will be crossed under double-masked conditions to the alternative treatment.

5.1.b. Sample: For the first stage, 400 depressed outpatients from 4 clinical sites will be randomized into one of the two treatment arms (CIT: 200, BUP: 200). Adults, age 18-65 will be included. Broad inclusion and minimal exclusion criteria (as in STAR*D) will ensure a highly representative and ecologically valid sample (see Section 8 for a detailed list of inclusion and exclusion criteria[3, 4]). All patients will meet DSM-IV TR[158] criteria for nonpsychotic MDD that is established using the SCID-I/P, to ensure a careful documentation of the phenotype for each patient (this includes recording and entering all symptoms of MDD for each patient), as well as documenting age of onset, number of episodes, and other manifestations of chronicity which have been associated with medication treatment response.[159]

5.1.c. Stage 1 Treatment Phase: Outcomes will be assessed in terms of response rates ($HRSD_{17}$) and tolerability (FIBSER) at Stage 1 exit (up to 8 weeks of treatment). Treatment will be guided by clinician-rated symptom measures (the 16-item Quick Inventory of Depressive Symptomatology-self-rated[160, 161] or $QIDS\text{-}SR_{16}$) and global side effects measures (the Frequency, Intensity, and Burden of Side Effects Rating[162] or FIBSER) obtained at each treatment visit, as in STAR*D. Medication treatment visits will occur at baseline and at weeks

1, 2, 4, 6, and 8 to ensure delivery of appropriate, vigorous, and tolerable pharmacotherapy. Those with unacceptable/intolerable side effects despite dose reduction may elect to enter Stage 2 Treatment Phase. Patients that are defined as nonresponders (< 50% improvement) will enter Stage 2 Treatment Phase.

5.1.d. Stage 2 Treatment Phase: Nonresponding patients will be crossed over, under double masked conditions. PBO nonresponders will receive CIT, and CIT nonresponders will receive BUP, Again, treatment will be guided by self-rated symptom measures (the QIDS-SR_{16}) and global side effects measures (FIBSER) obtained at each treatment visit. Visit frequency, dose escalation, and treatment monitoring will follow the same procedures used in Stage 1, response and remission will be defined as a 50% improvement and $HRSD_{17} \leq 7$, respectively.

5.1.e. Why include Stage 2: Including the Stage 2 treatment phase results in a design similar to a SMART design.[1] This SMART-like design is to cross over nonresponders to a treatment they did not receive in Stage 1 Consistent biosignatures in both phases may be suggestive of the strongest biologic correlates of treatment response.

5.1.f. Patient Flow (see Figure 2): Of the 200 patients randomized to CIT, about 90 will respond to treatment, and roughly 70 will be nonresponders (we estimate that the remaining 40 will be dropouts). For those randomized to PBO, we estimate about 60 responders, and 100 nonresponders. Patient flow estimates are based on our STAR*D experience, and from prior placebo-controlled antidepressant treatment trials.

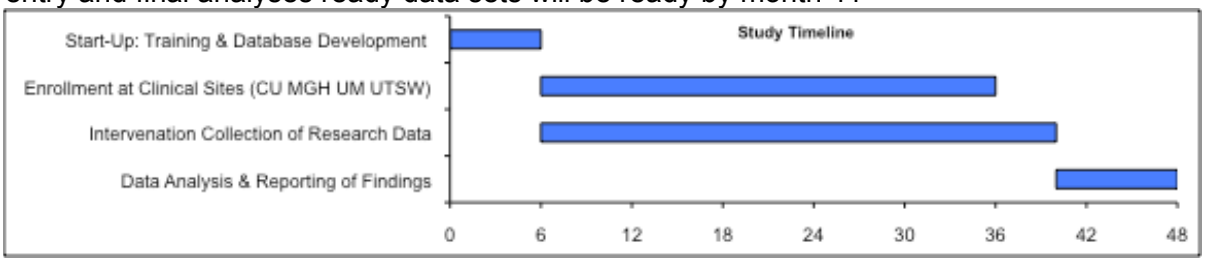
5.2. Study Timeline (see Figure 3)

5.2.a. Start-up (Months 1-6): During the first 6 months we will: (1) complete the Clinician and Clinical Research Coordinator (CRC) Procedures and the Data Procedures manuals; (2) establish the training and quality control procedures to ensure the delivery of protocol-driven high quality care; (3) program and complete pilot testing of the data acquisition and transfer protocol; (4) procure and distribute medications to the clinical sites (CSs); (5) hire additional clinical research coordinators (CRCs), and train all CRCs and Study Clinicians in all study procedures; (6) obtain final IRB approvals for the study; and (7) ensure completion of all ethics, human subjects, and HIPAA training for all key study personnel.

5.2.b. Enrollment and Intervention (Months 7-40): Study enrollment will begin at all 4 CSs immediately after start-up. We expect to enroll 400 participants over the ensuing 2.5 years (half of years 1, all of 2 and 3), roughly at a rate of 2.9 patients per month per site. Patients will enter a 16-week treatment period with Stage 1 Treatment Phase being completed in the first 8 weeks and Stage 2 Treatment Phase completed in the next 8 weeks.

5.2.c. Data Analysis and Reporting of Primary Analyses (Months 41-48): The last participant will enter Stage 1 Treatment Phase at month 36, and will complete Stage 2 Treatment Phase at month 40. All data entry and final analyses ready data sets will be ready by month 41 and proposed primary and secondary analysis will be completed by month 45. Final manuscripts reporting primary results will be submitted for publication by month 45.

Figure 3



5.2.d. Developing the DTRI to Personalize Treatment (Months 41-48): These 8 months will incorporate the most promising preliminary study findings towards the development of the DTRI, a clinically significant algorithm to personalize antidepressant treatment selection. A follow-up study to test the DTRI against standard clinical practice will be designed during this time. Also, data generated by this initiative will be broadly shared to drive the design of future validation studies.

5.3. Rationale for Durations of Stage Phases and Types of Treatments

5.3.a. Duration of Treatment Stage Phases 1 and 2: The duration of each Treatment Stage will be 8 weeks. While a longer trial would have advantages in allowing more time for patients to attain response and remission, the inclusion of a placebo comparator group makes a trial longer than eight weeks problematic from both ethical and practical perspectives.

5.3.b. Type of Treatment (CIT, PBO, BUP): The rationale for the inclusion of placebo is described above (*See 1.2.1)*; simply stated, having two treatments which are as different as possible from one another maximizes the ability to find a moderator and this study seeks to find moderators of antidepressant medication. The selected antidepressants are pharmacologically distinct active treatment arms: CIT (serotonergic) and

BUP (nonserotonergic), and provide an opportunity to explore moderators and mediators of differential treatment outcome for two active treatments.

5.3.c. Protocol for Treatments: To mimic clinical practice, enhance safety, ensure a vigorous dosing regimen, and maximize generalizability, all participants and treating clinicians will be able to titrate doses. A clinical treatment manual, with an emphasis on measurement-based care,[5] will recommend starting doses and dose changes on the basis of scores for the self-rated QIDS-SR_{16} and the FIBSER obtained at each treatment visit.

In addition, didactic instruction, supported by the clinical research coordinators (CRCs), and a centralized monitoring system with feedback will ensure timely increases in doses when an inadequate reduction in symptoms will occur in the context of acceptable side effects. Treatment will be aimed at symptom remission.

CIT *(stages 1 and 2)*: CIT will be started at 20 mg/day and then its dose will be raised to 40 mg/day by day 14 (week 2) and to 60 mg/day (final dose) by day 28 (week 4). The dosing of CIT will be the same in the two treatment phases (Stages 1 and 2).

PBO (stages 1 and 2): PBO will be started at 1 pill/day and then its dose will be increased to match the dosing instructions of the prior/other group.

BUP *(stage 2)*: The extended-release formulation of BUP will be started at 150 mg/day for seven days, 300 mg from day 8 to 27 (weeks 2 to 4), and 450 mg from day 28 (week 4) onward. The dosing of BUP is exclusively for CIT nonresponders to Stage 1 who are willing to switch to another active treatment.

5.4. Screening and Inclusion Criteria: Subjects will enter the study and be randomized to one of two treatments. This trial will be conducted according to the FDA guidelines and the Declaration of Helsinki. Written informed consent will be obtained from all patients before protocol-specified procedures are carried out. The subjects will be drawn primarily from an outpatient sample of patients with MDD, diagnosed by the use of the Structured Clinical Interview for DSM-IV Axis I Disorders - Patient Edition (SCID-I/P). At study entry, subjects must meet SCID criteria for a major depressive episode (MDE) and have a Montgomery-Asberg Depression Rating Scale (MADRS)[155] score of ≥20 at the screen visit. In addition, all patients must have not failed to respond to any prior trial of an antidepressant in the current episode at an adequate dose and duration, as defined by the MGH Antidepressant Treatment Response Questionnaire (MGH-ATRQ).[163] The MGH-ATRQ provides specific criteria for adequate dose and adequate length of a trial for it to be considered a failure. Patients who have shown inadequate response or poor tolerability to CIT or BUP in the past are not eligible (see inclusion and exclusion criteria in section 8).

5.5. Outcome Assessments: Clinical ratings are provided by three different staff members with distinct roles. The clinical rater-1 (CR-1) conducts the diagnostic interview (SCID-I/P) to determine diagnosis and study eligibility. This individual also completes the MADRS at screening. The blinded clinical rater−2 (CR-2) provides the primary measure of antidepressant efficacy ($HRSD_{17}$). The CR-1 and CR-2 are individuals with substantial clinical experience in this population. Although this is a double-blind study, to further protect against potential unblinding of placebo we feel it is important to obtain assessments of clinical changes in an independent fashion. As such every site will have a clinical rater (CR-2) who will rate outcomes, independent of any treatment knowledge. The first 3 months of Year 1 will be devoted to training staff and establishing inter-rater reliability across raters. Prior to enrolling the first subject, all raters will meet in Dallas to establish inter-site reliability. Training will be standardized under the direction of the study PIs, although we expect that all raters will have prior experience of at least 2 years in the use of the key rating instruments. While some rater turnover is expected over the course of the 4-year grant period, each site has sufficient infrastructure and redundancy of research operations to enable smooth transitions. The CR-1, CR-2, and SP are extensively trained in the conduct of interviews and the use of the relevant instruments, and participate in the study only after meeting certification criteria. The CR-1 will receive comprehensive classroom instruction provided by the PIs and watch the SCID 101 Didactic Videotape series developed by Dr. Michael First and Ms. Miriam Gibbon of New York State Psychiatric Institute. Classroom training will be followed by live observation and then supervised interviewing of patients, as well as generation of diagnoses from role-playing exercises. A focus of training is on establishing common thresholds for psychosis, hypomania (BPII), and schizoaffective disorder, since each condition is exclusionary. The CR-1 intake interviews are videotaped in a random 10% of patients screened for the study, including those rejected. These tapes are blindly reviewed by a senior diagnostician to provide on-going feedback regarding the interview quality and quantitative assessment of diagnostic reliability.

To conduct clinical ratings the CR-1, CR-2, and SP must be certified in the use of the $HRSD_{17}$, the primary clinical outcome measure and in the MADRS as well. The CR-1 rates the MADRS only at the pre-treatment baseline to ensure sufficient symptom severity to meet trial entry criterion. The CR-2 and SP make multiple longitudinal assessments at critical time points allowing for assessment of inter-rater reliability in the key outcome measures and the early detection of rater drift (CR-2's ratings are used for statistical analyses). For

purposes of certification, each rater completes the $HRSD_{17}$ and MADRS for 5 videotapes of patients with MDD that had been rated by 3 experts. Certification criteria have been developed for MADRS with thresholds for intraclass correlation coefficients (ICCs) for total MADRS scores (study rater vs. expert mean), median ICC for item scores, absolute discrepancy in total score, and average discrepancy in total score. To be certified, each rater must exceed threshold level of performance for each criterion. Certification failure results in one additional evaluation, after retraining and discussion with the Principal Investigators. Failure to pass a second evaluation means that the rater cannot contribute evaluations in this study.

5.5.a. Diagnostic Evaluation: The following instruments will be administered in the course of the study (see Table 2) to perform a comprehensive diagnostic evaluation, to assess clinical features, to gather demographic information, and to obtain information on prior treatment history during the current major depressive episode:

1) Structured Clinical Interview for DSM-IV: The SCID-I/P, administered by the CR-1, proceeds by modules to diagnose the different Axis I disorders. Questions here are asked exactly as written, and each is based on the individual criteria from DSM-IV. Answers are generally rated on a scale of 1-3 (1= doubtful, 2= probable, 3= definite), and based on the number of positive answers, diagnoses are determined. The SCID-I/P allows raters to gather demographic information, to assess psychiatric comorbidity, and to assess clinical features, including whether the subject meets criteria for atypical or melancholic major depressive episodes.

Table 2. Schedule of Clinician/Subject Clinical Ratings & Tests – Phases 1 and 2

Measurement	Screen / Baseline Day 0	Visit 1 Day 7	Visit 2 Day 14	Visit 3 Day 28	Visit 4 Day 42	Visit 5 Day 56	Visit 6 Day 70	Visit 7 Day 84	Visit 8 Day 112 Endpoint
SCID-I (CR), (SR) , FHS (CR), BHS (SR)	X								
NEO-FFI, SAPAS	X								X
ATRQ (SR), AAQ (SR) and ASRM (SR)	X								
CTQ (SR)	X								
RLCQ (SR)	X			X		X		X	X
SIGH-D_{24}	X	X	X	X	X	X	X	X	X
SIGMA	X								X
QIDS-SR	X	X	X	X	X	X	X	X	X
SAS (SR),	X	X	X	X	X	X	X	X	X
AEs, 15-item SAFTEE (SR), and SFI	X	X	X	X	X	X	X	X	X
CHRT (SR), CSSRS (CR)	X	X	X	X	X	X	X	X	X
CAST (SR)	X	X	X	X	X	X	X	X	X

2) The Standardised Assessment of Personality – Abbreviated Scale (SAPAS) will be administered to assess the presence of personality disorders.[164] The NEO-FFI (**Neuroticism-Extroversion-Openness Inventory-Five Factor Inventory**[165]**)** is the abbreviated assessment of the Five Factor Model of Personality (NEO-I).

5.5.b. Symptom Measures: The following instruments are administered throughout the course of the study (see Table 2): 1) the Structured Interview Guide for the Montgomery-Asberg Depression Rating Scale (MADRS)[155] (SIGMA); 2) the Structured Interview Guide for the Hamilton Rating Scale for Depression – 24 items (SIGH-D_{22})[156, 157, 166]; 3) the 16-item Quick Inventory of Depressive Symptomatology-self-rated (QIDS-SR_{16})[160, 161]; 4) the 17 item Concise Associated Symptoms Tracking (CAST); 5) the Columbia Suicide Severity Rating (C-SSR-C)[167]; 6) the 14-item Concise Health Risk Tracking (CHRT); 7) the spontaneously-reported Adverse Events (AEs) form; 8) the Childhood Trauma Questionnaire[168] (CTQ); 9) Self-Administered Comorbidity Questionnaire (SCQ)[169]; 10) the Family History Screen (FHS)[170]; 11) the Recent Life Changes Questionnaire (RLCQ)[173]; 12) The Altman Self-Rating Mania Scale (ASRM)[172]; 13) the Mood Disorders Questionnaire (MDQ)[179]; 14) the 15-item SAFTEE[174]; 15) the Sexual Functioning Inventory (SFI)[175]; and 16) the self-rated Anger Attacks Questionnaire (AAQ)[176, 177].

5.6. Measures for Clinical Management: *(symptom severity, treatment side effects, treatment adherence):* Treatment will be guided by self-rated symptom measures (the QIDS-SR_{16}) and side effects measures (the SAFTEE-SI) obtained at each treatment visit.

5.7. Biomarker Methods

5.7.a. Clinical Signatures: We will collect data on clinical phenotypes that have either been shown to predict treatment outcomes in MDD, or for which there is literature implicating them as being linked to poor outcomes. Clinical phenotyping will also serve as baseline assessment prior to randomization in the RCT.

a. Anxious depression: We will assess for anxiety both as a continuous (HRS-D Anxiety/Somatization Factor Score ≥7) and categorical variable (SCID-I Anxiety Disorders Section)

b. Chronic depression: We will use variables from the SCID Axis I Mood Disorders module to record the duration of the current episode, number of episodes, and the presence of DSM-IV "chronic" MDE (i.e., duration ≥ 2 years) to estimate disease chronicity. In addition we will profile chronicity using items from a CU chronicity form, used as a brief addendum to the SCID, which has been shown to be predictive of treatment outcome.[178]

c. Depressive subtypes: Subtypes of melancholic and atypical depression will be assessed both as a continuous variable ($HRSD_{29}$) and as a categorical variable (SCID- Axis I Major depressive Episode subtyping). The self-rated Anger Attacks Questionnaire (AAQ)[176, 177] will be used at the screening visit only to establish which MDD patients report anger attacks.

d. Trauma history: The Childhood Trauma Questionnaire (CTQ) is a 28-item self-report questionnaire used to assess childhood physical and sexual abuse and neglect.[168]

e. Family History: We will obtain family history of mood disorders in first degree relatives using the Family History Screen (FHS).[178] This ~15 minute interview screens for presence of major psychiatric conditions in first-degree relatives. It has been shown to be valid and reliable and is widely used.

f. Self Administered Comorbidity Questionnaire (SCQ)[178] will be completed at baseline by the patient to gauge the severity/morbidity of general medical conditions by organ systems, The SCQ will be gathered in addition to standard history and physical exam also performed at baseline.

g. Physical Symptoms of Depression (including insomnia, hypersomnia, fatigue, and psychomotor retardation): These symptoms will be extracted from the $HRSD_{29}$, which measures psychomotor retardation (and agitation), insomnia, hypersomnia, and fatigue.

h. Sociodemographics: we will use a Baseline Demographics Form to collect data on sex, race, ethnicity, employment status, marital status, years of education, and occupation.

i. Length of index episode: We will assess length of index episode based on the SCID I.

j. Concurrent psychiatric disorders: Because we will administer the complete SCID I, we will document the presence of other comorbidities including anxiety disorders and past history of alcohol and drug use disorders.

k. Subthreshold Bipolar Symptoms: The Mood Disorders Questionaire[179] is a self-rated form that will be used to capture the lifetime occurrence of manic and hypomanic symptoms in our participants with unipolar depression to explore the concept of bipolar spectrum. Baseline function: This is be assessed using the Social Assessment Scale (SAS),[180] which measures function in work, school, marriage, parenting, extended family, and leisure activities.

m. Current Life Stressors: Employment and marital status (single vs. married or cohabiting), and total number of life events at baseline assessed using the Recent Life Changes Questionnaire (RLCQ) by Miller and Rahe[173]. This is an update of the Holmes Rahe Scale and is a frequently cited and widely recognized life events checklist consisting of 87 items which takes 10-15 minutes to complete. We would modify it to suit the study time frame, 12 months at baseline, and 4 weeks at week 4 and week 8. This would let us look at stress as both a moderator and a mediator. Personality Disorder and Temperament: The NEO-FFI[165] a 60 item self report measure of the Five Factor Model of personality (Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness to Experience). The Standardised Assessment of Personality – Abbreviated Scale (SAPAS)[164] is a brief screen for personality disorder. A score of 3 on the screening interview correctly identified the presence of DSM–IV personality disorder in 80% of participants. The sensitivity and specificity were 0.94 and and 0.85 respectively.

o. Additional Clinical Assessments: Pain Frequency, Intensity and Burden Scale (P-FIBS). The P-FIBS is a 4-item self-report evaluating the frequency, intensity and burden of pain over the past week, as well as usage of pain medication to manage pain. SADS item 246 (modified) assesses past suicidal thoughts and behaviors (1 minute CR)

Table 3 details all baseline assessments and time required for clinical phenotyping. We also include those instruments that will be used for monitoring and as outcome measures in the RCT.

Table 3: Clinical Phenotyping and Monitoring in RCT (see Appendix B for instrument description)

Name of Instrument	Brief Description	Domain	Phenotyping		Monitoring
			Screening	Baseline	
SCID I	Axis I Diagnoses (clinician rated [CR])	MDD diagnosis MDD subtype	60 mins	--	--

		Comorbidities: anxiety and substance use			
NEO-FFI	60 item SR	5 Factor Personality Inventory	10 mins	--	--
ATRQ	Antidepressants and adequate treatment criteria	Past Antidepressant Treatment Response	5 mins	--	--
SAPAS	8 item SR	Personality disorder diagnoses	2 mins	--	--
SIGMA	10 item symptom CR	Depression-dimensional	10 mins	--	
SIGH-D	32 item CR	Depression-dimensional	--	10 mins	
QIDS-SR_{16}	16 item SR	Depression-dimensional	--	5 mins	10 mins
ASRM	5 item SR	Bipolar Symptoms	--	1 min	--
AAQ	7 item SR	Anger attacks-syndromal	1 min	--	--
CTQ	28 item SR	Childhood abuse and neglect	--	10 min	--
PERI-LES	15 item SR	Current Life Stressors	--	5 min	--
FHS	15 item CR	Family History of Depression	--	15 min	--
SCQ	15 item SR	Medical Comorbidities	5 mins	--	--
B-Demo	Basic sociodemo CR	Sociodemographics	10 mins	--	--
SAS-short	20 item SR	Social Adjustment	--	10 mins	10 mins
SAFTEE (modified)	15 item SR	Treatment emergent side effects	--	5 mins	5 mins
CHRT	14 item SR	Treatment emergent suicidality	--	2 mins	2 mins
CAST	17 item SR	Treatment emergent suicidality	--	2 mins	2 mins
C-SSRS	8 item CR	Treatment emergent suicidality	--	PRN	PRN
SFI	6 item SR	Treatment emergent sexual dysfunction	1 min	1 min	1 min
Patient burden			**95 mins**	**60 mins**	**60 mins**

5.7.b. Neuroimaging Methods

Neuroimaging assessments will be performed on all 400 participants prior to starting treatment and at one week after Stage 1 treatment initiation, (except DTI, which will be collected at baseline only). Pretreatment and early response neuroimaging measures will be examined for differential moderation of response to treatment by the end of the first 8-week, treatment period. All fMRI measures will be acquired in an identical fashion at pretreatment and week 1 for each participant, with task order counterbalanced across participants to avoid order effects.

General Imaging Methods: Neuroimaging data will be collected using a 3.0 Tesla MRI scanner at each site using a multichannel head coil, with consistency of image quality assured by a quality control procedure (see below). Each site is specifically selected because of their extensive expertise in MRI and functional imaging, past experience in multi-site investigations, and extensive physics support to match protocols. The protocol has four main components (1) high contrast / resolution structural imaging amenable to automated quantitative assessment of morphometric volumes and cortical thickness; (2) targeted task-based fMRI protocols designed to elicit activity in specific systems important in emotion processing and implicit emotion regulation, and reward processing, and (3) resting-state connectivity measures able to assess functional correlations in intrinsic activity across neural systems, and resting state PASL to assess regional cerebral blood flow; (4) Diffusion tensor imaging (DTI) that will allow for fractional anisotropy and tractography measures. Stimuli will be presented on back-projection screens attached to the head coil, and subjects will respond using a standardized button box.

Pre-scanning: Subjects will be given instructions needed for each task and the opportunity to practice prior to entering the scanner. Performance of >90% accuracy on practice trials (where relevant) will indicate that the subject learned the task, finger response mapping and stimulus timing to criterion.

Scan step 1: Structural 3D axial MPRAGE images will be acquired (TE: 3.29 ms, TR: 2200 ms, Flip angle 9°, Field of view: 256x192 mm, Slice thickness: 1 mm, Matrix: 256x256, 192 continuous slices, 7:02 min).

Scan step 2: Subjects will be given four fMRI tasks (emotional conflict, reward processing, PASL, and resting-state connectivity acquisition). A brief practice run will be re-presented to each participant (and performance verified) prior to the functional scan to ensure adequate performance. Acquisition parameters for T2*-weighted (BOLD) functional images will be: 39 axial slices (3.1mm thick, TR/TE=2000/28msec, FOV=205x205cm, matrix=64x64; Flip angle90°). Run length will be 13:14min (emotional conflict), 8:02min (reward), two, 6 min resting state, scans, and a 5 min 28sec PASL sequence. For the resting state and PASL scans, subjects will fixate a fixation cross.

Scan step 3: There will also be a 10 min DTI data (diffusion weighted images –DWI) acquisition. This will be performed using echo planar imaging (voxel size: 2 x 2 x 2 mm, 61 no-colinear directions; b value: 1000 s/mm2 and seven additional volumes with no-diffusion weighting that will be acquired distributed throughout the acquisition sequence.

The order of the above tasks will be: structural MRI acquisition, PASL, first resting state acquisition, emotional conflict /reward task, reward task/emotional conflict task, DTI, second resting state acquisition. The order of the two fMRI tasks will be counterbalanced across participants.

5.7.b.1. Activation-based task-specific methods:

1) Emotion processing and implicit emotion regulation: Emotional conflict task. This task will be performed as previously described[181-183], are asked to identify the expression of a face (fearful or happy) while ignoring an overlying emotion word (“fear” or “happy”). Emotional words thus either match (congruent) or conflict (incongruent) with the facial expression. Stimuli are in a pseudo-random order, counterbalanced across trial types for expression, word, response button, and gender. Behavioral data include reaction times for correct trials, accuracy, and post-error adjustments in reaction time and accuracy.

2) Reward processing: Reward Task. This is a slow event-related card-guessing game[184] that allows examination of neural activity during anticipation and receipt of monetary reward feedback. Each trial is a possible win, where outcome is either win or no-change (disappointment), or a possible loss, where outcome is either loss or no-change (relief). Trials are presented in pseudorandom order with predetermined outcomes. Participants are told that their performance will determine a monetary reward after the scan, with $1 for each win and 50 cents deducted for each loss (total $3). During each trial, participants guess the value of a visually presented card (4s), learn the trial type (possible-win or possible-loss), anticipate feedback (actual win, disappointment, actual loss, relief−6s), and receive outcome feedback (1s). Behavioral data are reaction time.

Behavioral data are:

1) Emotional conflict task: reaction times for correct trials, accuracy and post-error adjustments in reaction time and accuracy;

2) Reward task: reaction times during card value guessing (as participants anticipate possible reward or loss).

Data Processing

fMRI data analyses: fMRI data generated from each task will be analyzed and preprocessed in Pittsburgh, led by Dr. Phillips’ team, with support from Dr. Etkin. The key dependent variables include BOLD waveform index of activity in the following regions of interest for each task and specific stimulus contrasts:

1. Emotional Conflict Task*:* in amygdala and orbitomedial prefrontal cortex (OMPFC), including ventral prefrontal cortex, anterior cingulate gyrus (ACG) and mediodorsal prefrontal cortex (mdPFC) for post-congruent incongruent trials (i.e. trials higher in conflict due to less regulation) and post-incongruent incongruent trials (i.e. trials lower in conflict due to greater regulation);

2. Reward Task: in ventral striatum and OMPFC for anticipation of win, anticipation of loss, and outcome (win, loss, disappointment, relief).

We will also examine effective and functional connectivity between these key regions, and resting state connectivity. The details on the analyzing and processing are presented in detail in Appendix II.

3. Resting state analysis: Resting state connectivity data will be analyzed in Pittsburgh, led by Dr. Phillips’ team, with support from Dr. Sheline. We will use several complimentary analytic approaches on low-pass-filtered fMRI data (<0.1Hz). In seed-based analyses we will compute voxelwise correlations across the brain with time courses extracted from ROIs (e.g., OMPFC, amygdala, ventral striatum), controlling for global signal and motion parameters as regressors of no interest. This method is optimal for distinguishing between connectivity of adjacent regions (e.g. amygdalar subregions) or small structures (e.g. ventral striatum). Test-retest reliability of such measures has been in the moderate to high range for data acquisition lengths comparable to those acquired here (for extensive review, see Van Dijk et al[185]). We will also conduct several network analyses that assay global characteristics of brain architecture (see appendix). These include

graphical analysis, multivariate pattern classification using support vector machines, and independent components analysis. These methods are more ideally suited for understanding the operation of multiple, interconnected regions as coherent networks. Such measures have been recently published by grant investigators to capture cortical integration[393] and specialization of brain systems.[186]

4. Perfusion Imaging (Arterial Spin Labeling) analyses. Perfusion weighted image series will be generated by pair-wise subtraction of the label and control functional images, followed by conversion to absolute cerebral blood flow (CBF) image series. Clusters showing associations between CBF and treatment response will be identified for the whole brain at a significance level of $p<0.005$ (uncorrected) and cluster size larger than 50 voxels, as in previous perfusion studies.

5. DTI Analysis. Done at CU. Our pipeline for processing DTI data safeguards against artifacts before diffusion-weighted images (DWIs) are reconstructed to diffusion tensor maps by 1) visually inspecting all DWIs for motion or susceptibility artifacts; 2) correcting distortions along the phase-encoding direction that are induced by eddy currents[187] 3) co-registering DWIs within each subject using a rigid body transformation using ANTS; & 4) correcting nonlinear geometric distortions in the DWIs using intensity-based diffeomorphic registration to a ANTS template[188] or by DTI-TK.[189] We then reconstruct a diffusion tensor at each voxel and compute tensor statistics, including fractional anisotropy as well as deterministic and probabilistic tractography using FSL (FMRIB) and Camino (http://www.cs.ucl.ac.uk/research/medic/camino/).

6. Structural MRI Analysis. Done at CU. Brain extraction and cortical thickness will be done in one of two ways, to be selected on the basis of an ongoing evaluation. One approach will make use of standard Freesurfer algorithms and the second approach will make use of two new algorithms, Atropos and DiReCT, both part of ANTS (http://www.picsl.upenn.edu/ANTS/). FreeSurfer's (http://surfer.nmr.mgh.harvard.edu/) Hybrid Watershed Algorithm[190] extracts brain matter using white-matter connectivity and a deformable surface model. Atropos uses nonlinear registration to initialize a GMM-MRF multi-class brain extraction model; the model uses geometric constraints on topology and tissue-connectivity to isolate the brain from extraneous structures. FreeSurfer measures cortical thickness[117] by generating accurate models of both the gray/white and pial surfaces and measuring the distance between these two surfaces. DiReCT measures cortical thickness by computing a continuous one-to-one correspondence (diffeomorphic mapping) between gray/white matter boundaries and gray matter/cerebrospinal fluid boundaries, and measuring the distance between these boundaries.

7. Behavioral data analyses for the specific reaction time measures described above for each task will be analyzed using parametric and nonparametric tests as appropriate in SPSS.

5.7.b.2. MRI Quality control: Quality control will involve three distinct goals (1) to ensure sequences are collected appropriately at each site (e.g., to ensure the pulse sequences were executed as specified in the protocols), (2) to monitor between-site factors that affect data comparability such as scanner SNR and particular scanner geometry properties and drift, and (3) to vet the each data session for artifacts or quality problems (scanner noise) and medical problems.

To ensure sequences are collected appropriately, all data will be uploaded into a centralized XNAT repository hosted that will parse DICOM header information for required field information. The XNAT system, which is presently being used for large multi-investigator studies, was developed by Randy Buckner and Daniel Marcus.[191, 192] In addition to header parsing, automated pipelines presently exist to assess subject movement during functional paradigms and estimate SNR in BOLD-images. As data are uploaded, these procedures will be run providing rapid feedback about protocol deviations. To ensure cross-site reliability of neuroimaging data, we will obtain data from identical MRI phantom acquisitions monthly from each of the neuroimaging sites. Well established routines for using phantoms, (objects with known MRI properties when scanned), will be employed to perform quality assurance on the scanners used in this study. This approach will allow consistency of signal-to-noise across scanners over the time, and to homogenize activation in our neural regions of interest in each task across the different sites will allow us to control by covarying for inter-site differences in signal to noise to homogenize activation in our neural regions of interest in each task across the different sites, as recommended (e.g., Friedman et al.[193, 194]). This approach will also minimize other variabilities in image acquisition and quality between the sites, ultimately ensuring that consistent data will be obtained at each site. We will adopt the data acquisition and information sharing standards published by the Biomedical Informatics Research Network (BIRN), detailed at http://www.nbirn.net/. To vet individual data sets for image artifacts and medical problems, each image session will be viewed directly by a trained research assistant to look for evidence of spiking, image distortion, aberrant head position, etc. The above mentioned quantitative estimates will be used to monitor

functional protocol SNR. Each image session will be examined by a board certified neuroradiologist (e.g., stroke, lacunar infarcts).

Scanner and RF coils. The same scanner strength will be used (*3T*). We propose to converge towards using the same multi-channel RF coil to reduce inter-site variability at the level of data acquisition. Images will be acquired either with an 8-channel coil or a 12-channel coil. For those sites that do not have an 8-channel RF coil, images will be acquired with a 12-channel RF coil and will then be made comparable with images acquired on the *default 8-channel RF coil* by comparing their reference RF field maps in order to adjust their intensity based on a correction factor. The *correction factor* will be derived from the intensity ratio between images acquired with the two different coils. All sites will optimize their sequence to achieve the most favorable SNR and CNR profiles. However, across platforms on different manufacturers we will require comparable SNR and CNR (within 10% of each other).

Diffusion Tension Imaging (DTI). DTI data will be acquired across 4 different sites with different types of scanners. Our goal is to reduce inter-site variability by carefully calibrating the acquisition of data across different sites. This calibration will be done by using an isotropic phantom and subsequently the same human phantom across all sites. We plan to take similar measures as those used for calibrating DTI data in the BIRN DTI project, proposed and carried out by Dr. Susumu Mori's group at Johns Hopkins University.
We plan to use the DTI protocol to collect 4 sets of DTI data on the same human *phantom* within 24 hours. We will then investigate the effects of the following parameters on the reproducibility of DTI-derived data: signal-to-noise ratio (SNR) and number of gradient orientations. Our data will be acquired at b=1000 s/mm^2 along 25 gradient directions plus 3 baseline images. From this dataset, we will generate DTI-based contrast using data from 1 dataset (lowest SNR) up to 3 datasets (highest SNR). Also, by using subsets of the 25 orientations, we will study the impact of the number of gradient orientations. For example, we may extract data with 10 orientations and signal average 3 times (total 25 diffusion-weighted images). We can compare this data with 1 set of 25 orientation data. We can also measure the effect of the number of b_0 images. Based on these evaluations, we will finally either verify or revise the suggested DTI sequence. If further inter-site calibration is necessary, we plan to acquire data at different TE values, in order to study the impact of this parameter on the reproducibility of DTI data. Because a b-value of 1000 will be consistently used, we will not study the effect of differing b-values.

In order to measure the reproducibility of various analysis results within each site, the entire procedure will be repeated in 3 sessions at different occasions using the same subject. This intra-site calibration should be performed only after we have agreed upon a DTI sequence. However, to simplify the procedure and save cost, each site can choose to study the reproducibility individually using their *local human subject* to run this 3-session evaluation, because the intra-site test-retest reliability is a local issue.

Eddy currents can be corrected using appropriate post-acquisition image processing techniques. Distortions induced by eddy currents along directions other than the phase-encoding direction can be ignored, and an iterative cross-correlation algorithm referenced to CSF-suppressed (such as FLAIR) images, or calibrations from water phantoms, can be used to correct accurately eddy current distortion in images with b-values as high as 2000 s/mm^2, fully sufficient for implementation in our studies at 3T. Also, because short echo train lengths are used to acquire DWIs, the b_0-related DWIs have minimal distortion. Therefore, these b_0-related DWIs can be used as the undistorted baseline against which the remaining DWIs are nonlinearly coregistered, thereby correcting any nonlinear geometric distortions.

Each time a DTI dataset is acquired, a research assistant (RA) should be present to verify that all the key imaging parameters are correct, including but not limited to the b-value, slice orientation, head position, TR/TE, excitation number, matrix size, slice thickness and gap, field of view, head coverage, and proper attachment and positioning of a marker to the subject's head. Because a DTI sequence is relatively lengthy, motion artifacts are most likely to happen. *Therefore, we propose that the RA either monitor the entire scanning procedure of the DTI data acquisition, or that the RA manually go over all the DTI data right after the data are acquired, to ensure there is no visually apparent motion. When motion is visually detectable, the DTI scan should be repeated.* The same ongoing quality control system should also address and monitor other major problems, such as excessively low SNR and abnormal distortion. A brief note should be recorded for any unusual observation. Otherwise, "pass" should be stamped for passing the ongoing quality check. *Finally, we should clarify that we do not plan to scan the DTI phantom after each subject scan (this is different than what is proposed for the other modalities)*.

Structural Imaging *In order to evaluate image quality and reliability across different sites, we propose to acquire images using both a phantom and a human subject. We agree to use an Agar phantom (for calibration of SNR) as well as a structured phantom which has markers inside placed in a rectilinear fashion to help assess for geometric distortions (we will use the structured phantom available at NYSPI).* The same phantoms and human subject will be used for each of the sites to assess the quality of imaging data acquired across these sites. The same pulse sequence must be employed to acquire both the human and the phantom data. This pulse sequence will be the 3D MP-RAGE sequence developed for our 3T GE scanner (NYSPI), optimized for the specific scanner at each institution. We will review the scan data to ensure that the parameters are correct and to maximize the scan quality. For the phantom scan, phantom positioning inside the coil should be reproducible across sites. Therefore, there should be visible external markers to aid in this positioning. The center of the phantom should be aligned with the center of the coil. One may use the alignment lights on the scanner to position the phantom into the center of the magnet. A localizer will be obtained to ensure that the phantom is placed correctly. Scan data will be acquired left-to-right in sagittal orientation.

For *human subject scans*, accurate positioning is also essential (at all stages of this study). The subject should be positioned in the same manner for each and every scan, placing a stereotactic marker on the left side of the subject head. The centering crosshair of the scanner should be aligned on the subject's nasion (between the eyebrows) at every scan and the head coil should be centered over the subject's head, making sure that the subject is deep enough in the coil to prevent signal loss at the inferior aspect of the brain. The subject should be advanced to the isocenter of the scanning bore and his or her head should not be rotated but centered laterally along the inter-hemispheric fissure. Proper placement in the head coil is crucial because scans are acquired straight, not in an oblique orientation. Full brain coverage will be necessary because studies that do not contain the whole brain and skull cannot be processed**.** The skull must be fully included superiorly and laterally. The entire cerebellum should be included inferiorly. In the anterior -posterior plane, the nose should also be included otherwise image folding (wrap around artifact) will result. Anatomical data will be acquired as 1 NEX, with two scans per subject acquired one after the other.

To test reproducibility of the quality of images acquired at each site, the entire procedure will be repeated 3 sessions at different occasions (e.g. 1 per day) using a human subject (different than the subject used for inter-site calibration). Because reproducibility is a local issue, real-subject data can be acquired using a subject selected by each site independently.

Minimizing image variances in images caused by differing imaging platform is essential for quantitative analysis. For example, nonlinearities in imaging gradients introduce distortions in images acquired from different sites, thereby limiting the validity and reliability of the quantitative analyses. Typical distortions in anatomical images can be classified as (1) geometric distortions, and (2) intensity nonuniformities.

Linearity in imaging gradients is usually a trade-off to improve the performance characteristics of the gradients. However, nonlinearities in the gradients introduce geometric (or spatial) distortions in the images. Differing imaging systems, which are available commercially, have differing built-in software to reduce spatial distortions. For example, in Philips systems, the acquired images are sufficiently linear, and therefore do not usually require any correction for spatial distortions. In Siemens scanners, software is available for correcting 2D in-plane distortions, (Some Siemens scanners have software for correcting distortions in 3D) However, these corrections may not be the default mode on specific Siemens scanners. Finally, in GE Systems, correcting spatial distortions in 2D (i.e., within slice correction) is the default-mode operation of the scanner, although software for correcting 3D nonlinearity is also available. At this time it is not possible to standardize these correction methods across scanners, therefore we suggest that each site use the correction method available for their platform. Furthermore, images acquired on all systems have residual spatial distortions because of the inaccuracies in the software algorithms available on the scanners. To further minimize residual spatial distortions in our images, we will therefore use the method for correcting gradient nonlinearity that is available from (BIRN) [http://www.nbirn.net/downloads/gradient_non_linearity/index.shtm]. This method uses the spherical harmonics expansion of the gradient coils used in the MRI system to generate a table of x, y, and z displacement of each voxel within the imaged volume. This table is then used to spatially displace the voxels in the image to correct for the spatial distortion across the image.

We will correct nonuniformities in image intensity caused by nonhomogeneous characteristics of the RF coil designs (B_1 inhomogeneities) by using an automated method called N3 (nonparametric nonuniform intensity normalization). These nonuniformities cause blurring of the image intensities, increasing image entropy and N3 restores image intensities by decreasing the entropy automatically.

The effects of distortion correction on the reproducibility of data across different sites as well as on the effects of field strength and manufacturer type on the stability of morphometric variables has been studied. We will continue to consult with experts at BIRN to ascertain what solutions they have implemented to minimize this variability based on their data.

Preprocessing nonuniformities in image intensity will be corrected by N3 (above). Extra-cerebral tissues are removed using an automated tool for extracting the brain (Atropos, see below).

To ensure scanner stability and scan quality throughout the study, each site should perform on going quality control scans on the structured phantom and on the Agar phantom immediately following each subject scan. Acquisitions with major motion artifacts should not be accepted and a repeat scan should be performed. Once the optimal sequence has been decided, good SNR must be assured. Because low SNR is usually caused by incorrect sequence parameters, a repeat scan with correct parameters will be required if low SNR becomes an issue. Finally, care should be taken to avoid signal losses caused by incorrect positioning of the subject in the head coil (for example, if the subject is placed too deep into the coil, a signal loss in superior regions may occur). If one site upgrades a scanner, NYSPI will be informed. To ensure overall consistency of the data, the site will be asked to perform a number of test scans on the phantoms and on a human subject.

5.7.c. Electrophysiology Methods

The sites in this project have established electrophysiology labs for measuring EEG and evoked potentials. Recording parameters will be standardized across laboratories to facilitate exporting and merging of data using Polyrex software developed at Columbia.[81] Dr. Tenke will travel to each site to ensure that stimulus delivery and EEG recordings are comparable across sites, and data will be sent to Columbia for quality control and merging of data.

Resting EEG. EEG will be recorded during four 2-minute periods, half with eyes-closed (C) and half eyes-open (O) in a counterbalanced order (COOC or OCCO). During the eyes-open condition, the subject fixates on a central fixation mark and tries to avoid blinking. During the eyes-closed condition, subjects are instructed to avoid eye movements.

Loudness dependency of N1/P2. The LDAEP paradigm requires subjects to sit quietly with their eyes opened and fixed on a cross while tones (1000 Hz, 40 ms duration with 10 ms rise and decay time) are presented at varying levels of intensity. Binaural tones are presented via headphones in a pseudorandomized order at five intensities (60, 70, 80, 90, 100 dB SPL) with interstimulus intervals ranging from 1600 – 2100 ms. Each stimulus intensity is repeated 100 times (i.e., 5 stimulus intensities = 500 trials).

EEG data acquisition and artifact procedures. Continuous EEG data will be recorded from at least 31 channels with a nose reference (for details see Bruder et al.[80]). EEG/evoked potential artifact procedures will include several proven screening and reduction routines, which have successfully been used to optimize EEG and evoked potential epochs. Artifact detection will also be verified by visual inspection. For spectral analyses, the blink-corrected continuous data will be segmented off-line into 1-s epochs every 0.5-s (50% overlap). For evoked-potential analyses, recording epochs of 1200 ms including a 200 ms prestimulus baseline will be extracted off-line. EEG spectra and ERP waveforms will be averaged from artifact-free data. ERPs will be low-pass filtered at 12.5 Hz (-24 dB/octave) and finally baseline-corrected using the 200 ms preceding stimulus onset.

Current source density (CSD). All artifact-free EEG epochs used for spectral analysis and all averaged ERP waveforms at each electrode will be transformed into reference-free CSD estimates using parameters previously established for a 31-channel recording montage.[195] By eliminating volume-conducted contributions from distant regions, CSD topographies have more sharply localized peaks than scalp potential topographies. Accurate and reliable surface Laplacian estimates have been obtained for low-density 31-channel ERP recordings.[83]

Data reduction. The averaged CSD spectra and CSD waveforms will be submitted to frequency (spectra) or temporal (waveforms) principal component analysis (PCA) derived from the covariance matrix, followed by unrestricted Varimax rotation of the covariance loadings.[81, 83, 84, 195] This approach determines common sources of variance in the reference-free transformations of the original EEG or evoked potential data in the form of distinctive PCA components (factor loadings) and corresponding weighting coefficients (factor scores). The spectral pattern or time course and topography of the extracted orthogonal factors allow identification and measurement of physiologically-relevant CSD components for further analysis.[82, 83, 195, 196]

EEG source localization. Previous work has implicated the theta band in predicting antidepressant response.[48, 49, 197] Thus, intracerebral sources (current density) of theta (6.5-8 Hz) activity will be computed

using Low Resolution Electromagnetic Tomography (LORETA[198]). The artifact-free segmented data used for scalp power analyses will be subjected to LORETA analyses following procedures developed by Dr. Pizzagalli's laboratory.[49] From scalp-recorded electrical potential distribution, LORETA computes the 3-dimensional intracerebral distributions of current density for specified EEG frequency bands. As in Pizzagalli et al,[49] we will use the LORETA version that includes a 3-shell spherical head model registered to the Talairach brain atlas,[199] as well as EEG electrode coordinates derived from cross-registrations between spherical and realistic head geometry.[200] Computations will be restricted to cortical gray matter and the hippocampus by using the digitized Talairach and Montreal Neurologic Institute probability atlases. The analyses will consist of 5 steps. First, for every participant, all available artifact-free EEG epochs will be subjected to cross-spectrum analysis for the theta EEG band (6.5-8.0 Hz). Second, LORETA will be used to compute current density as the linear, weighted sum of the scalp electrical potentials, which will be then squared to yield power of current density. Third, for every participant, the LORETA solution will be normalized to a total power of 1 and log-transformed. Fourth, using a region-of-interest approach, theta current density will be extracted from the rACC cluster previously found to predict treatment response,[49] and averaged across voxels. Fifth, the extracted value will be used to contrast eventual responders and nonresponders.

5.7.d. Behavioral Phenotyping Methods

Choice Reaction Time (5 min). The Choice RT task will be adapted from Thorne et al.[201] Subjects are presented with a black screen, on which four squares outlined in white are laid out in a windowpane pattern. A red "X" then appears in one of the boxes, and subjects must hit one of four keys a keyboard, the one that corresponds to the box that the "X" is in. After a response, the "X" disappears and, after a 50 msec delay, reappears in either the same or another box. Subjects are instructed to "catch the X" continuously by hitting the correct corresponding key. An incorrect response clears the X in each case, but is recorded as incorrect. A total of 60 items will be presented. Median RT is computed from all correct individual response times. A total correct response score is also computed.

Word Fluency (5 min). We will use the standard Controlled Oral Word Association test of Benton, using the standard letters F, A, and S.[202] In this task, subjects are required to generate words beginning with each of the three letters (one letter at a time) as fast as they can within one minute. The total number of words generated is summed.

A Not B Task (10 min): The A Not B task[141] is an adaptation of a paper-and-pencil working memory and reasoning task developed by Baddeley.[203] In this task, subjects are presented with a statement describing the arrangement of two letters (e.g., "A comes before B"). Below the statement are the two letters arranged as "AB" or "BA". The subject's task is to determine whether the statement accurately describes the arrangement of the letters, and to evaluate it as true or false (via keypress) as quickly possible. All permutations of the descriptions of the letters (A before B, B before A, A after B, B after A), as well as the negation of each of these permutations (e.g. A does not come before B) are paired with each of the two arrangements of the letters (16 total permutations, each repeated twice). Scores are the total number of correct responses (of 32 total items) and the median RT for correct responses.

Flanker Task (11 min). Subjects will complete a modified version of the flanker task. Each trial will consist of a vertical row of either congruent (>>>>> or <<<<<) or incongruent (>><>> or <<><<) arrows against a gray background. To increase task difficulty and induce more errors, which will be important for computing post-error behavioral adjustments, the flanking arrows will be presented first without the center arrow (to induce a prepotent response), and then the center arrow will be flashed. All arrows will then disappear and the participant will see a blank screen as they respond to the direction of the center arrow. Trials will involve the following sequence: fixation cross (490 ms), flanking arrows (80 ms), center stimulus plus flanking arrows (30 ms), interstimulus interval (1,500 ms). Subjects will perform two blocks, each consisting of 152 trials (block duration: 5 min, 20 sec). In order to increase the number of errors, in each block, 98 trials will be congruent (64.5%) and 54 trials will be incongruent (35.5%). RT and accuracy measures will be collected throughout the task. Primary analyses will focus on behavioral adjustments after error commission. Post-error adjustments[204, 205] will be operationalized as [Rabbitt effect = ($RT_{\text{After incorrect trials}} - RT_{\text{After correct trials}}$)] and [Laming effect = ($\text{Accuracy}_{\text{After incorrect trials}} - \text{Accuracy}_{\text{After correct trials}}$)], with higher scores indicating more adaptive behavioral adjustments. Based on our prior studies,[49, 137] a larger Laming effect is hypothesized to predict better treatment to citalopram but not bupropion. Secondary analyses will evaluate general task performance (e.g., overall RT and accuracy) as well as Flanker interference effects, which will be calculated as: [$RT_{\text{Incongruent trials}} - RT_{\text{Congruent trials}}$] and [$\text{Accuracy}_{\text{Congruent trials}} - \text{Accuracy}_{\text{Incongruent trials}}$], with higher scores indicating increased interference.

Probabilistic Reward Task (15 min). The task, which has been previously validated in several samples (e.g., ,[41, 148, 149] allows for the objective assessment of the subject's propensity to modulate behavior as a function of prior reinforcements. Briefly, in each trial, the subjects' goal is to determine, via button press, whether a short (11.5 mm) or a long (13 mm) mouth is presented on a previously mouthless cartoon face. The task will include two blocks, each involving 80 trials. Within each block an equal number of short and long mouths (n = 40) will be presented for 100 ms each. To elicit a response bias, an asymmetric reinforcer ratio will be utilized. Specifically, correct identification of either the short or long mouth will be rewarded ("Correct!! You won 20 Cents") three times more frequently ("rich stimulus"; n = 24)) than correct identification of the other mouth ("lean stimulus"; n = 8). The reinforcement allocation and key presses will be counterbalanced across subjects. Subjects will be informed at the outset that that their purpose will be to win as much money as possible. The main variable of interest will be response bias, which will be calculated using signal-detection theory, and indexes the systematic preference for the response paired with the more frequent reward ("rich stimulus"), or the extent to which behavior is modulated by reinforcement history. A high response bias emerges when subjects show high rates of correct identification for the rich stimulus and high miss rates for the stimulus associated with less frequent rewards.

5.8. Methods to Enhance Participant Adherence to Research Procedures: Participants will be paid up to $675 dollars for completion of all the research procedures. This includes $150 for clinical phenotyping (SCID, etc.), $100 for each MRI session (2 total), $50 for each EEG/neuropsychology testing session (2 total), $25 for screening blood sampling and $50 for genetic blood sampling, $50 for completion of follow-up ratings, and up to $50 per participant for necessary travel expenses. We will post a bimonthly (every 2 months) publicly available study update on the study Web site that Clinical Site Directors and CRCs can download for posting or distribution in the clinics. This update will include tips to CRCs and patients on how to enhance retention and adherence. At the end of year 1, certificates will be provided (no cost to the study) to recognize those CRCs and Clinical Sites with the highest enrollment, retention, and treatment adherence numbers.

5.9. Random Treatment Assignment Procedures: Randomization will be implemented via StudyManager software using a secure browser interface to the CUMC-IT. Each site will independently randomize patients accepted into the study using block randomization with random block sizes of 2, 3, or 4. This procedure will result in equal or nearly equal sample sizes in all groups within each site and obsfucate attempts to predict subject assignments to either arm.

5.10. Data Management & Processing: The DMC is directed by Phil Adams, Ph.D. The overriding aim of this group is to ensure data integrity while providing a efficient method of capturing all study data and managing data flow from multiple collection, processing and analysis sites. We will exercise strict control over the data stream at all times. All data releases (even within the group) will be regulated and recorded so that we know the status and location of specific data on all study subjects at all times. This status is centrally located in StudyManager (SM) and all data input and release(s) are recorded in SM. Study procedures and policies will state that the only permissible data flow s are receipts from the DataCenter (DC) or flows into the DC because external flows outside of this loop cannot be tracked or controlled.

The DMC includes two functional units. The StudyManager (SM) software will be used for electronic data capture and clinical trial management and will be hosted by Columbia University Medical Center/ Information Technology (CUMC/IT). The nonclinical data (imaging and EEG data) will be stored by MIND (The Molecular Imaging and Neuropathology Division at the New York Psychiatric Institute, NYSPI, see below). The DMC is responsible for storing all the data (clinical and nonclinical) and managing the flow of data from the 4 collection sites to centralized storage in the DMC, distributing raw data for processing to specific sites (see below for details) and receiving return flows/data from these sites. The DMC will have the responsibility of merging clinical and nonclinical data into files for analysis as specified by the Data Analysis workgroup. The SM system will be configured to track all incoming and outgoing datafiles in order to produce integrated reports with the potential to track every procedure conducted as part of the study. The DMC will also manage/track DNA samples as they are collected and sent to the Rutgers University Cell-DNA Repository. All operations of the DMC will be overseen by National Coordinating Center (UTSW) and coordination between these centers will be facilitated by an "Operations Conference Call" at frequent /weekly intervals.

Data Management & Processing can be broken into five phases: Study setup, Recruitment, RCT treatment and follow-up, Data Analysis/Publications, and post-study Data Sharing. The Data plan includes operations for each phase. Each of the four sites participating in the RCT is called a Collection Site. All collected data will use common study identifiers generated by SM. Nonclinical data (with the exception of DNA) will flow from the four Collection sites to MIND where it will be channeled to specialized Processing sites as follows: Structural MRI

and DTI will be processed at CU, fMRI quality control at MGH (Buckner lab), fMRI at Pittsburgh, EEG (pre-processing and scalp analyses) at NYSPI (Bruder lab), EEG (source localization/LORETA) at McLean Hospital (Pizzagalli lab), behavioral phenotyping at NYSPI (Bruder lab: Choice Reaction Time, Word Fluency, A Not B Task) and McLean Hospital (Pizzagalli lab: Flanker Task, Probabilistic Reward Task). After processing, the data will sent back to MIND and the SM system will be updated. Finally, clinical data from SM will be exported to MIND where it will be integrated for analysis and eventual distribution to a Data Sharing site (Rutgers University).

5.11. Summary of Data Management Operations across Phases

Study setup: Setting up SM to track pre-study calibration and standardization of procedures for nonclinical (EEG and MRI) data collection including checking machine calibration at (yet to be determined) intervals during the RCT. Definition and implementation of Recruitment tracking in SM including procedures, forms, schedules, staff access and tracking of eligibility/disqualification. Designing the StudyManager (SM) screens for collecting clinical data (electronic data capture) and design of the RCT and centralized randomization (if required by the design). Setup of all nonclinical assessments in SM with status codes to track processing. Setting up procedures and database for tracking file collection and processing of nonclinical data (MIND). (The granularity of tracking the data processing of raw, nonclinical data needs to be determined.) Implementation of a system to update SM automatically from the MIND database as data are collected and processed. SM needs to be setup to create invoices for subcontractors/sites. Definition, design and implementation of SM reports as required by the Executive Committee, NIMH, Study Coordination Center, Collection sites and the Data Management sites. Reports will cover the full array of all operations including Recruitment, Subject Retention, Compliance, Safety/Adverse events, etc.

Recruitment: All subjects who are presented an Initial Screening Assessment (with the potential for study exclusion) will be logged into StudyManager so we have an accurate record of the complete subject pool. As subjects progress through the trial, SM will maintain a single “Status” field for each subject that identifies the overall status of the subject in the study. In addition, all study procedures will be logged with their own Status and Date fields.

RCT treatment and follow-up: SM has the ability to track and inform the Study Coordination Team of progress toward study milestones. Data collection at sites requires logging procedures into SM and SM will be programmed to monitor the flow and processing of nonclinical data in a timely manner as prescribed by the Study Coordination Center. Monitoring adverse effects is also incorporated into SM.

Data Analysis/Publications: Data exports from SM are part of the systems’ Clinical Data Management Software (CDMS) component. MIND will develop scripts to integrate clinical data with nonclinical summary variables for analysis. Data sets will be archived with date-time stamps and distribution logged into SM so that we have a defined path of data distribution with the potential for both updating data sets and retrieving earlier distributions.

Data Sharing: In conjunction with NIMH we will develop a plan to share data from the RCT with the research community.

5.12. Clinical Trial Quality Control

Manual of Operations: The Manual of Operations (MO) will be a reference and training manual, that will be prepared during the planning phase to document general procedures (e.g., study organization, ancillary study procedures, publications policy, industry collaboration guidelines). In addition, the MO will include information on the protocol (e.g., treatment protocol, follow-up visit schedule), describe the details of the data collection and data transfer procedures, forms and question-by-question instructions for each data collection form, adverse event and serious adverse event procedures, and procedures for the collection and shipment of biological specimens. Copies of the manuals of operations will be bound in 3-ring binders and mailed to each investigator, as well as posted on the private area of the study web site.

Training, Certification, and Quality Control: Near the end of the start-up phase of the study, a training session will be conducted. The goal of this initial training will be to familiarize the investigators and research staff at each Clinical Site and laboratory on the procedures needed to conduct of the study. Material covered will include orientation to the study web site (e.g., explaining how to use the help desk), review of procedures, review of the MO, and description of the general data entry and management principals. Training sessions will be recorded and available on the study web site for investigators to review or to train new investigators. The training program, materials and certification will be developed by the DMC in conjunction with the NCC.

Data Entry and Management Training: The DMC staff will collaborate with the NCC to train Clinical Site personnel in all required procedures related to data entry and management. Training will include an overview

of the data entry and management system as well as a hands-on practicum. Certification will be awarded upon the attendance (or viewing of the recorded session on the study web site) and satisfactory completion of a short, web-based test given after training.

Intervention Training and Quality Control: Pharmacotherapists will need to attend an in-person training session or review the recorded training session on the study website. After completing the training session, pharmacotherapists will need to complete a certification exam prior to treating any study patients. In addition, the DMC will work with the NCC to develop a manual for protocol and reports to monitor adherence to the treatment protocol. Only licensed clinicians (based on local, state, or federal regulations) that have completed NIH human subjects training will be eligible for certification.

Retraining and Staff Turnover Training: During the course of the study, it will be necessary to re-certify staff to ensure that no subtle changes occur over time. Retraining will occur at annual meetings, and will include the review of study procedures and changes to study procedures as well as the completion of a re-certification exam. In addition, all trained staff who change duties and all new staff must be certified for their new duties. This can be accomplished by viewing training sessions on the web site and completing certification exams.

Reports: A number of reports will be generated to provide information on the status of each CORC protocol. Below is a summary of three broad types of reports that may be generated.

Study Monitoring Reports: Several types of reports will be provided to study investigators on various aspects the protocol through a protected area of the study web site. The available reports include static reports that are updated on a routine schedule, dynamic reports that are generated from a database at the time of request, and user-configurable reports that are customized to manage the data in the database.

Reports available via the Data Management System (e.g., the form status report, the duplicate status report and the case status report) are useful when troubleshooting database issues, documenting the successful completion of data processes, and provide views of the database from differing perspectives. Other reports will be used to summarize the progress, including recruitment and retention, schedule of upcoming interviews, treatment adherence and summary of data quality.

Steering Committee Reports: For each Steering Committee meeting, a comprehensive set of reports will be presented, including: 1) recruitment and follow-up, 2) quality control reports, 3) participant follow-up adherence data (including missed visits), and 4) participant characteristics.

DSMB Reports: In addition, the DSMB will review and evaluate study progress including data on recruitment, quality control, safety and follow-up retention. The DMC will work to ensure all of the information presented at the DSMB meetings will remain confidential. Each report will be marked as such. No interim treatment outcome information will be presented to study investigators or to the scientific community. At the conclusion of each DSMB meeting, all reports will be collected and destroyed.

We recognize that designing and implementing an ambitious project like EMBARC will require a thoughtful, nimble organization and operation that will pay close attention to a clinical trial while simultaneously obtaining a vast array of high quality biological markers in addition to clinical markers. Furthermore, the use of biomarkers has to be understood in the context of implementing a complex clinical trial if the results are going to be clinically meaningful. The specific organization and approach to conducting this study has been chosen to maximize our achievement of the study aims.

1) Study Organizational Structure – This section includes the overall structure of the study organization: the executive committee, the National Coordinating Center (NCC), National Data Management Center (DMC), Clinical Sites (CSs), and Scientific Cores.
2) Project Administration – This section provides a discussion of how the structure and associated functions will achieve the study aims and objectives. This section also addresses ways in which communication and coordination as well as data entry and data management will occur.

An infrastructure model which has been very successful in coordinating prior multi-center trials (STAR*D, CO-MED, TMAP) will be used for this study. It includes a National Coordinating Center, National Data Management Center and Clinical Sites. This infrastructure is enhanced for this trial by the addition of scientific cores, which will integrate the contribution of scientists who are experts in the biomarkers that will be collected and analyzed.

5.13. Study Organizational Structure: The PIs (Drs. Trivedi and Weissmann) oversee the scientific coordination of the study and manage the entire experienced infrastructure, with the help of the NCC, DMC and the scientific oversight of the scientific cores. The infrastructure also includes the research delivery system of 4 Clinical sites where the study is conducted and clinical and biological markers are collected, and Drs. Kurian and McGrath (Psychopharmacology), who oversee the delivery of study interventions and ensure high quality training and implementation. Drs. Trivedi and Weissman will also oversee the functioning of the 3

Scientific Cores, each consisting of scientific experts in the collection, storage, development and analyses of the specific biomarkers that will be collected—Clinical, Imaging, and Neurophysiology. The scientific integrity and safety of the study is overseen by the NIMH Data and Safety Monitoring Board (DSMB).

5.13.a. NCC Coordinating Center: The NCC, under the leadership of Dr. Trivedi, is a hub of the infrastructure. The PI, with the help of Dr. Maurizio Fava and the other NCC leadership is responsible and has final accountability for the scientific, clinical, operational, administrative, and financial oversight of the study as well as the publication and dissemination of methods and findings. The PI and NCC under the guidance of the Executive Committee provide or coordinate clinical oversight, training, and quality control to ensure successful conduct of the study (including implementation of the psychopharmacology) and collection of all study clinical and biological markers. Backgrounds and functions for all NCC individuals are described in the budget justification. The NCC carries out its functions with guidance from the Steering Committee, which recommends policies for training, implementation, and management of the study, as well as for the publication and dissemination of study procedures and findings.

The NCC includes Madhukar Trivedi, M.D., (PI and NCC Director), Benji Kurian, M.D. (psychopharmacology oversight and safety monitor), To Be Named (project and performance management), Andrew Kozel, M.D. (fMRI Coordinator), David Morris, Ph.D. (clinical assessment training/control), and administrative staff.

5.13.b. Data Management Center (DMC): see above (Section 5.10)

5.13.c. The Research Delivery System and Scientific Cores:

5.13.c.1. Clinical sites (CSs):

All the four Clinical Site Directors (CSD) for this study are experienced with both STAR*D and CO-MED. As a group, we have repeatedly demonstrated that we can enroll very large samples of participants within expected timeframes. Details of the track record of clinical sites are in Section 4.1 of the Research Plan.

CRCs (who will also function as clinical raters) are available at all sites (one primary CRC and one CRC who will provide blinded ratings) as well as resources for blood sample collection and handling and collection of neuroimaging and neurophysiology markers.

CSDs are responsible for the scientific and operational performance of the trial at their respective CSs. The CSDs are responsible for ensuring that recruitment, retention, treatment fidelity, and clinical data collection targets for assessments and all biological samples and images are met. The CSDs ensure local Institutional Review Board (IRB) approvals, human subjects training certifications, and conformance to all regulatory requirements (e.g., ethics, HIPAA) for the RC. The CSDs identify and oversee that CRCs, the study physicians that deliver the pharmacotherapy carry out their responsibilities (while Dr. Kurian the quality of the delivery of psychopharmacology). The CSDs will develop and implement corrective actions should any performance not meet standards and requirements

The CRCs are responsible (along with the study physicians) for the management of patients and the study protocol at each CS. They collect intake and clinic visit information, manage and transmit patient data, and ensure all assessments and biomarkers are collected.

The CSDs train the study physicians in all protocol and operational requirements, with assistance from the PI and the pharmacology experts in the Clinical Core.

5.13.c.2. Scientific Cores (SCs): There are three Scientific Cores: Clinical, Imaging, and Neurophysiology, each of which has a core leader as well as co-leaders for the specialty areas within the core. Each of the Scientific Cores also each meet as a group, including all of the scientists with expertise in the core's area. These experts are responsible, coordinated by each core's leader, for the content of the training of the CSDs, CRCs, study physicians, and technicians collecting data in their respective areas, implementation of the collection, transmission and analysis of marker data, and quality control. The core leaders are responsible for coordinating the efforts of the other experts and consultants in their respective areas who will be contributing to this project in the areas of training and implementation of the study. The core teams meet as a team twice a month during study start-up and monthly until the end of the study. Core leaders, specialty area leaders within each core, and other experts in each scientific core will be central to the analysis and interpretation of study data and publications in their respective areas of expertise. The Core leaders and specialty area leaders are included in Table 4 below.

In addition to training, specific responsibilities are as follows:

In the Clinical Core, Dr. Benji Kurian will provide day-to-day oversight of study physicians' fidelity to the psychopharmacology protocol, using measurement based care to make treatment decisions. Dr. David Morris will supervise training, quality control, and day-to-day oversight for the collection of clinical assessments.

In the Imaging Core, fMRI expert Mary Phillips, M.D. (University of Pittsburgh) will oversee acquisition, processing and analysis of fMRI data. Furthermore, Dr. Phillips will supervise the 3 tasks performed during imaging and oversee quality control for each fMRI task. Additionally, Dr. Phillips will oversee that Amit Etkin, M.D. will process and analyze data for the facial recognition task, and Yvette Sheline, M.D. will consult with Dr. Phillips to process and analyze resting connectivity data. Imaging quality control core expert Randy Buckner, Ph.D. will develop standard parameters for the imaging equipment and procedures, while also overseeing quality control of the physics of the fMRI task on an ongoing basis throughout the trial. In addition, Dr. Buckner will work with imaging technicians, and CRCs, in running periodic phantom experiments.

Similarly, in the Imaging Core, Ramin Parsey, M.D., Ph.D. will oversee acquisition, processing and analysis of the structural MRI data. This includes DTI analysis (fractional anisotropy, deterministic and probabilistic tractography) and cortical thickness measures. Dr. Parsey will work with Dr. Kangarlu in ensuring the quality of DTI and structural data. He will work with Dr. Klein in the analysis of the DTI data.

In the Electrophysiology and Behavioral Phenotyping Core, Gerard Bruder, Ph.D., will lead the administration and acquisition of qEEG data and will be assisted by Diego Pizzagalli, Ph.D. Dr. Bruder will oversee pre-processing of the scalp EEG data, scalp power analyses, and analyses of the LDAEP data. Dr. Pizzagalli will oversee source localization analyses of both EEG and LDAEP data. Both Dr. Pizzagalli and Dr. Bruder will share responsibility for overseeing the behavioral phenotyping tests.

All raw data collected at each of the CSs will be forwarded to the DMC, who will then forward the data to the specific Scientific Core for processing of raw data, which will in-turn be sent back to the DMC for statistical analysis.

5.13.d. Data Safety and Monitoring Board (DSMB): The National Institute of Mental Health's (NIMH) DSMB meets three times a year to consider and review protocols and consent documents for multicenter Clinical trials sponsored or administered by the NIMH. The DSMB monitors safety issues, including the review of adverse events; the adequacy and integrity of accumulating data; and the study's capabilities to test its hypotheses. DSMBs approve study initiation and beginning subject enrollment, and determine if study procedures should be altered or stopped because of evidence of benefit or harm to trial subjects that may be attributable to one of the interventions under evaluation or reasons related to scientific integrity. DSMBs conduct independent and objective reviews of all accumulated data, and consider requests to conduct interim analyses. The NCC responds to all DSMB issues and queries and ensures that all DSMB actions and correspondences are submitted to local IRBs throughout the trial.

5.13.e. Committee and Core Team Structure: The committee structure builds on the core committee and administrative structure used effectively for ten years in STAR*D and CO-MED to meet or exceed study performance targets (STAR*D enrolled more than its planned 4000 participants and CO-MED completed and exceeded planned enrollment before its aggressive targeted enrollment completion date). We have also generated over 100 publications of findings and methods from STAR*D and the Depression Trials Network to date. Committees include the Steering Committee, NCC Study Management Team, Operations Committee, and Publications Committee, all used productively for years.

Meetings of each Scientific Core are one of the key mechanisms by which the experts in each Core meet and organize and co-ordinate their activities from start-up through study implementation. Each will include both Dr. Trivedi and Dr. Weissman, and all of the experts contributing to this study in each specific area. The Cores will also work together on analysis of data and publications. Core meetings will be led by the Core Leader and Co-Leader. Madhukar Trivedi, M.D. and Myrna Weissman will be members of all Cores to ensure that all functions of each Core are well coordinated with all study activities. The NCC Study Management Team will assist the Cores as needed. Memberships of each Core are in Table 4. Minutes for all meetings will be documented and disseminated.

5.13.e.1. Steering Committee (SC): The Steering Committee meets weekly and serves as the executive decision making body for the study. It approves the scientific, clinical and administrative policies and procedures, makes major policy decisions as the study is implemented, and reviews and implements recommendations from NIMH and the DSMB. It oversees that the study is conducted and monitored effectively, and is making progress in achieving its aims and that all co-investigators are functioning effectively. It also makes decisions on key operational issues; reviews the effectiveness of data collection, and interpretation; and

oversees the performance of CSs on performance measures (e.g., fidelity to the protocol, recruitment, retention, outcome data collection) and oversees that all core experts are performing their functions effectively and in a coordinated manner. Representation from all scientific Core Teams, the NCC, DMC, and Clinical Sites will be included.

5.13.e.2. NCC Study Management Team*:* The NCC Study Management Team (M. Trivedi, M.D., Chair) meets weekly and is responsible for the day-to-day functioning of the study at the CSs and DMC. It implements policies, procedures and takes action to address day-to-day scientific, clinical, operational and financial management issues to enhance the efficiency and performance of the study based on weekly and monthly performance reports, and other sources of information. It also follows up to monitor improvements. This team oversees the adequate delivery of pharmacotherapy and assists the Core Teams. At the discretion of the Steering Committee and/or PI, it also follows up to ensure that any identified issues in the CSs or scientific cores are resolved.

5.13.e.3. Operations Committee: The Operations Committee (P. Adams, Ph.D., Chair) meets weekly during the study to manage operational issues related to the ongoing conduct of the trial at the CSs, with a focus on data collection and integrity, such as monitoring the overall quality of study data, implementing needed changes to the protocol and/or data collection forms, overseeing site visits and audits, and certification and recertification for study personnel. This committee also monitors that the quality control activities of the cores are being performed.

5.13.e.4. Publications Committee*:* The Publications Committee (**M. Fava, M.D., Chair; P.J. McGrath Co-Chair; MH Trivedi, MM Weissman, R Parsey and M Phillips**) meets monthly or semi-monthly (as of month 18) and oversees the development, writing and submission of research publications, chapters and other scientific communications. The committee develops plans for publications, and develops the focus and designs the data analytic approach for the core set of publications. It oversees appropriate authorship for manuscripts and presentations, timely analysis of findings and timely completion of manuscripts following data closure, the scientific integrity and quality of the manuscripts, and the completion of final publication. All publications are reviewed by the Publications Committee before submission. This committee also considers data requests from other investigators outside this project. Finally, there will be opportunities both in the first two years of the study, as well as after data collection has been completed, to add ancillary studies. The committee will encourage the field to participate in ancillary studies, evaluate proposed ancillary studies for inclusion, and oversee the publication process for any ancillary studies that are included. This committee has successfully overseen over 100 publications from STAR*D, CO-MED and other Depression Trials Network studies.

The first task of the Publications Committee will be to develop a publications plan and publications policy for the study. The Committee will include specific time requirements in the Publication plan and policy (e.g., first draft expected within two weeks of the completion of the data analyses). Once developed, the plan and policy will be circulated to all research staff and posted on the web site. The study will generate several types of manuscripts. Procedural manuscripts will focus on issues related to the design and conduct of the study (e.g., Design and Rationale). Other manuscripts will focus on data collected and include the analyses of the primary and secondary outcomes. This approach has been successfully used in the STAR*D and Co-Med trials

5.13.e.5. Imaging Core (M. Phillips, chair, R. Parsey, co-chair), Neurophysiology Core (Gerry Bruder, chair, D. Pizzagalli, co-chair) and Clinical Markers Core (M. Fava, chair, P. McGrath, co-chair): The Cores will meet twice a month during study start-up and the first three months of study implementation and monthly thereafter. Each Core will produce study policy and procedures in its respective area, develop all training plans and materials, oversee quality control activities where applicable, oversee ongoing data transmission and specialty analyses of data prior to forwarding to the DMC (e.g. EEG, imaging) and all other core issues related to the conduct of the trial. These Cores will be key to the interpretation of study findings and production of publications for the field.

Table 4 Study Committees and Cores

Steering Committee (meets weekly) – *Members to Be Named*			
Publications Committee (meets monthly) – *Members to Be Named*	DMC Operations Committee (meets weekly/twice a	NCC Study Management Team (meets weekly) –	Scientific Cores (meets twice a month/monthly) –

	month) – *Members to Be Named*	*Members to Be Named*	*Members to Be Named*

5.13.f. Project Administration: The PIs and NCC, in collaboration with the DMC, Committees, Cores and CSDs, oversee that all activities necessary to achieving the aims of the study are successfully implemented and that all deliverables are met within specified time frames. Communications and study management occur using the Core teams and committees, and other study management teleconferences.

5.13.f.1. Scientific Oversight, Clinical Oversight, Training, Quality Control, and Study Operations by the NCC and DMC: The PIs provides scientific leadership, direction and oversight for all aspects of the study. This accountability is woven through virtually all study activities from the development and management of the study infrastructure and protocol to the development and execution of all clinical, research and data procedures, any needed protocol changes and the dissemination of methods and findings through publication and other means. The PI is present on virtually all committees. All NCC leadership, the DMC and CS leaderships report to the PI. The PI also oversees the activities of all the core experts via the core teams and NCC Study Management Team. The committees and teleconferences bring together the scientific investigators and operational leaders throughout the infrastructure in a highly organized manner, proven historically to be effective and efficient, for input regarding start-up and execution of the study, and analysis and publication of findings.

5.13.f.2. Development of Clinician and CRC Procedures Manual: The NCC, DMC, and Scientific Core Team experts, collaboratively develop standardized study procedures and materials including data collection forms, and formalize them in a Clinician and CRC Procedures Manual.

The NCC has final accountability for these documents and procedures. The manuals include an overview of the study aims and detailed descriptions of the protocol, inclusion/exclusion criteria, general principles and specific issues in treatment implementation, medication dosing, side effects management, serious adverse event reporting, an overview of personnel responsibilities, patient screening and intake procedures, obtaining informed consent, instructions for patient access to study medications, clinic visit procedures, monitoring of fidelity to the protocol, data collection and transfer procedures, including the provision of clear definitions and instructions for completing the data forms, and interpreting the results , as well as detailed instructions for the collection and transmission of all biological samples, images, and EEG data. Copies of the manuals of operations will be bound in 3-ring binders and mailed to each investigator, as well as posted on the private area of the study Web site.

5.13.f.3. Development and implementation of initial and ongoing training for study personnel: Clinical and research procedure trainings are conducted by NCC and scientific core experts identified for each target area and audience. Many of the CSDs and CRCs, and NCC and DMC personnel, are experienced with running other studies, including STAR*D and CO-MED, and are expected to implement this study smoothly. We have conducted the psychopharmacology protocol with measurement based care, and blood sample collection in other studies for years. What this study adds is collection of a neurophysiology and imaging markers for all participants. We are confident these can be easily integrated into study procedures by our experienced team. All training sessions will be video recorded and stored on the web site for future reference.

Training and Certification: Comprehensive training occurs in Dallas for CSDs, CRCs who function as clinical raters (one unblinded, one blinded) and the CRC who will coordinate fMRI collection and transmission during the first “All Hands” meeting. Training covers study aims and protocol, issues in the protection of human subjects and IRB processes, clinical treatment and research procedures, including fidelity to the protocol and safety procedures, research procedures including understanding reliability and validity, data collection and transmission and other data integrity procedures, administration of study assessments, collection, processing and transmission of biological samples, and activities required to collect and transmit qEEG and fMRI data. Extensive description of the assessment training and certification of the two CRCs who will function as clinical raters at each site and the study physicians in the conduct of study assessments is in Section 5.5 of the Research Plan. Certification will be provided for the fMRI CRCs as well. The DMC staff will collaborate with the NCC to train Clinical Site personnel in all required procedures related to data entry and management. Training will include an overview of the data entry and management system as well as a hands-on practicum. In addition, specific training will be held for the data managers. Certification in data management will be awarded upon attendance at training or viewing of a recorded training session in data management on the website and satisfactory completion of a short web-based test.

Intervention Training: Psychopharmacologists will attend an in-person training session with CSDs, utilizing the Clinician and CRC Procedures Manual, and a training outline and materials developed at the NCC under Dr. Trivedi's direction, or review the recorded training session on the website. After completing the training session, psychopharmacologists will need to complete a certification exam prior the treating any study patients. In addition, the DMC and NCC will develop a manual for protocol and reports to monitor adherence to the treatment protocol.

Staff Turnover Training: In addition, all trained staff who change duties and all new staff must be certified for their new duties. Assessment, protocol, handling of biological markers/data, and data entry training and certification will be provided for new staff as they join the study.

5.13.f.4. Development and implementation of quality control for data procedures, psychopharmacology: Data quality control procedures, including point of entry quality control, are developed and implemented by the DMC. Effective implementation of clinical procedures is supported by thorough training of CRCs, CSDs, study physicians, and CRCs in clinical procedures and quality control mechanisms including ongoing monitoring of adherence to the protocol (for psychopharmacology).

The Clinician and CRC Procedures Manual defines the psychopharmacology treatment protocol – when and whether a treatment, dosage, or type should be changed based on symptoms and side effects. Simple, easy-to-use measures that define symptoms and side effects are used to inform treatment decision-making and to enhance appropriate fidelity to protocol treatment recommendations.[160-162]

A comprehensive flagging system using summary reports helps enhance fidelity to the protocol for psychopharmacology treatments. As data on medication dosage, global burden of side effects and symptoms are entered by the CRC during each patient visit, protocol deviations will be flagged. The CRC can, therefore, review medication dosing issues with the clinician. A deviation may be justified or may be a management error needing adjustment. The DMC generates detailed weekly reports by CSs, and study physicians for CSD review during weekly meetings with the study physicians and CRCs, and for NCC review.

Training updates in clinical and research procedures and special topics are reviewed on an ongoing basis during scheduled and ad hoc teleconferences (CRCs, CSDs, study physicians). Training updates are addressed during the annual "All Hands" meetings in Dallas, for CSDs and study physicians on site visits by Dr. Trivedi and other NCC staff. Comprehensive adherence activities are completed on an ongoing basis by Dr. Kurian for all study physicians.

Site visits and audits are performed by the Clinical Manager/Assessment Manager to identify any clinical, protocol related, data, or regulatory issues requiring addressing.

During the course of the study, it will be necessary to re-certify staff to ensure that no subtle changes occur over time. Retraining for CRCs will occur at the annual "All Hands" meetings, and will include the review of study procedures and changes to study procedures. Inter-rater reliability, based on standardized videotaped clinical interviews and gold standard scores, will be established annually.

See Section 5.7 of the Research Plan for a comprehensive description of quality control for biological markers.

5.13.f.5. Development, implementation and oversight of safety procedures: CRCs and study physicians provide a net of safety procedures to identify suicidal risk. As in any clinical setting, the clinician treating each patient may be the first to identify that the patient is at risk for suicide during a contact. Study clinicians may hospitalize patients, increase visit frequency, change treatments within the protocol, or remove the patient from the study if indicated. The CRC also evaluates the patient at each visit, has an ongoing relationship, and passes any indication of risk immediately to the study physician or clinician for action.

A general description of a serious adverse event (SAE) is an unanticipated event involving risk to study participants. The reporting of SAEs will be responsibility of each Clinical Site. The SAE procedures for this study will be similar to procedures used in other studies, such as CO-MED. The procedure will be developed so that SAEs are processed within 24 hours and evaluated to determine whether or not an SAE meets expedited reporting criteria to the Food and Drug Administration (FDA). For this study, the procedure will be initiated when a Clinical Site investigator enters a SAE form into the study database. The system would then automatically sends e-mail notification (with Participant ID and event date, and a link to the SAE report describing the event) and a SAE Summary form to Dr. Kurian, safety officer at the NCC, who is responsible for processing SAEs and if decided upon in advance, representatives of the NIMH. The Safety Officer completes a series of questions via the web-based system. Using these data, an algorithm determines whether the SAE meets the requirements for expedited reporting to the FDA. Depending upon the review, the files are moved to

the appropriate area of the web site (Expedited SAE or Nonexpedited SAE), and e-mail notifications are sent to the Clinical Site and NIMH. A full report of SAEs is provided in the DSMB reports that are to be submitted at least three times per year, based on current NIMH policies. Dr. Kurian also provides recommendations for any needed follow-up to the patient's study physician and/or CS Director. The SAE reporting procedures are detailed in the Human Subjects Section.

5.13.f.6. Development and implementation of performance management and quality improvement for the conduct of the clinical trial: The overall functioning of the infrastructure, as well as its performance and efficiency in the study, is maximized by using an organized quality improvement system, effectively used to maximize performance and efficiency in STAR*D and CO-MED. Specific critical targets chosen for clinical and administrative efforts include patient screening, recruitment, retention, successful collection of research and clinical data, as well as measures of fidelity to protocol treatment. The system effectively identifies areas that require intervention to maximize performance. Systematic reporting of all identified goals is developed and monitored (at least monthly) by the NCC Study Management Team. Once processes responsible for lagging performance are identified, a specific plan and timeframe for correction action is formulated and implemented by the appropriate group or individual (e.g., CRC, Clinical Manager) under the supervision of the NCC Study Management Team. Corrective actions, along with a revisiting of barriers and revisions of the corrective action plan, continue until the performance indicators meet the predetermined goal.[206]

5.13.f.7. Development and implementation of a comprehensive study teleconference/ communications strategy: For communications in such a large and complex study to run effectively and efficiently, an organized, comprehensive, communications system is needed. In addition to teleconferences for the committees and cores with national membership (all of them) there are ongoing teleconferences for the CSDs and study physicians (weekly), led by Dr. Kurian, and CRCs (weekly), led by the clinical manager.

For this study, a hub of the communications system will be the web site developed and managed by the NCC, based on a combination of common web technologies, Microsoft SharePoint technology, streaming audio and video teleconferencing software and services. It is our intention that through the use of this system, the study will be provided with immediate communication, effective collaboration and efficient project management. The model will be the integrated technology based communications system, used in STAR*D and CO-MED.

The web site will be comprised of two main areas, an area open to the ***public*** and a ***private*** area for researchers. The information contained on the ***public*** web site will be targeted for the general public, those interested in the treatment of depression. The aim of the ***private*** area of the web site is an intranet designed to enhance the communication and organization among the cooperating institutions. Access to the Intranet will be restricted to investigators, research staff and committee members (e.g., Operations Committee).

The web site involves individual web panels, each representing a single feature or tool (e.g., email, web-based data entry, help desk, calendar, etc.), which will vary in presentation and availability based on a defined user's role (e.g., investigator, coordinator) and group association(s) (e.g., Steering Committee member).

A number of tools will be available on the Intranet. A brief description of these features follows:

Directory: A directory that allows personnel to update their own contact information will be present.

Calendar: A calendar system will be included where events (e.g., Steering Committee conference call) may be scheduled or modified by all study personnel. The calendar web panel, by default, will list the next several meetings or events that are relevant to the current user. In addition, a link to the complete calendar of events will be available.

Committee Resources: Committee resources will be accessible from a single web panel and will list committees (e.g., Publications or groups (e.g., Clinical Site Principal Investigators), based on a user's membership. The committee area is designed to organize and provide easy access to information (e.g., mission statements, meeting agendas, minutes) by committee or group. In addition, specialized tracking systems may be developed to aid committees. For example, to assist the Publications Committee on the oversight of the publication process, a web-based Publications Monitoring System (P&P) may be used to monitor the progress of manuscripts through to publication.

Document Library: Document sharing areas will be available for the sharing of manuscripts, operations memos, data collection forms, manuals of operation and training materials.

Reports: Reports will be routinely posted to a specialized reports section

Study Update: A study Update will be posted weekly. This will generate an e-mail to all study personnel. In other studies that we have coordinated, we found these weekly updates to be a very effective tool for keeping

all study personnel informed of project activities. The Update will briefly summarize all study activities during the course of the past week and will provide reminders for any deadlines or any upcoming activities in the next week. All Updates will be sequentially numbered to permit the study staff to easily store and review archived Updates. In addition, all Updates will remain on the web site to be reviewed at any time by all study personnel. This method of communication may also be used to notify personnel at the clinical sites of a protocol change or update, such as revisions to questionnaires, clarifications of questions or revisions in various study procedures. If changes are made to a form or a procedure, the change will be described and information will be provided on effective dates and location of documentation on the web site (e.g., new version of the data collection form).

Help Desk: A Help Desk/Frequently Asked Questions (FAQ) web panel will be used to assist study staff at the clinical sites or core laboratories when attempting to obtain answers to questions. The Help Desk web panel will list all unresolved help desk requests submitted by the researcher and the most frequently accessed FAQ entries. A search dialog will permit the researcher to find other entries based on keywords. If the FAQ database does not sufficiently aid in finding an answer to an issue, a direct link to the Help request form may be selected. The Help request form is to be completed and submitted online. An e-mail notification to the most qualified person (e.g., data manager) to address a particular issue based on the category or description of the help request is sent. Once a request is resolved, the individual will receive an e-mail verifying the solution.

Multi-Media Conferencing and Collaboration: The NCC and DMC have experience utilizing web meeting services such as AT&T Web Meeting Service and Microsoft Live Meeting. For this study, one can envision using such technologies for a number of tasks including the collaboration on study protocol design, manuscript preparation, re-certification of trained staff and the certification of replacement staff.

5.13.f.8. Reports for study management: The DMC will provide reports to assist in managing the study. Below is a summary of three broad types of reports that will be generated.

Study Monitoring Reports and Tracking Systems: Several types of reports will be provided to investigators on various aspects of the protocol through a protected area of the web site. The available reports include static reports that are updated on a routine schedule, dynamic reports that are generated from a database at the time of request, and user-configurable reports that are customized to manage the data in the database. Other reports will be used to summarize the progress of the protocol, including recruitment and retention, a schedule of upcoming interviews, treatment protocol adherence, and summaries of the quality of the data and procedures (e.g., percent missing data). Other specialized reports may be developed as well.

Steering Committee Reports: For each Steering Committee meeting, a comprehensive set of reports will be presented, including: 1) recruitment and follow-up, 2) quality control reports which include the quality of data received, 3) participant follow-up adherence data which include missed visits, and 4) participant characteristics.

DSMB Reports: The DSMB will review and evaluate study progress including data on recruitment, quality control, safety and follow-up retention. The DMC will work to ensure all of the information presented at the DSMB meetings will remain confidential. Each report will be marked as such. No interim treatment outcome information will be presented to investigators or to the scientific community. At the conclusion of each DSMB meeting, all reports will be collected and destroyed.

5.13.f.9. Development of consents and acquisition of initial and ongoing regulatory approvals: The NCC Regulatory Affairs Manager, Dr. Kathy Shores-Wilson, with over thirteen years of experience managing regulatory affairs for STAR*D, CO-MED and other trials, oversees the development of final consent forms by the senior NCC leadership for forwarding to CSs, tracks initial IRB approvals, and site activations of each CS when regulatory requirements are met. CSDs are responsible for ensuring that IRB, human subjects, and HIPAA training requirements and documentation are completed for study staff, physicians and clinicians.

5.13.f.10. Day-to-Day Oversight of Study Operations: Clear accountabilities for the full scope of functions performed during the study, is essential. The organizational structure has a clear chain of command. CSDs have direct supervisory responsibility for the performance of their CS and are overseen by the NCC. The scientific cores are accountable for the procedures, quality control, and initial evaluation of marker data in their areas of expertise but the PI has final accountability for all study activities.

5.13.f.11. Resolution of disputes: As in prior studies, it is expected that most or all disputes are resolved through consensus, including differences of opinion regarding site improvement procedures and reallocation of resources; however, clear lines of authority for decision making and dispute resolution are present. Study performance expectations of the DMC, CSs, and scientific Core experts are outlined in their respective subcontracts and include adequate delivery of performance goals, identified by the Steering Committee at the initiation of or during the study. Final accountabilities for addressing of disputes are defined in the contracts.

5.13.f.12. Provide Budget and Grant Management: The NCC provides budget and fiscal management for the study. The CSs and subcontracting organizations manage their own budgets and staff. The NCC provides payments for patients using the database effectively used for patient payments in CO-MED. We are very experienced in overseeing very large numbers of subcontractors. For STAR*D and CO-MED we managed over 60 subcontracts, handling all invoicing and payment related functions.

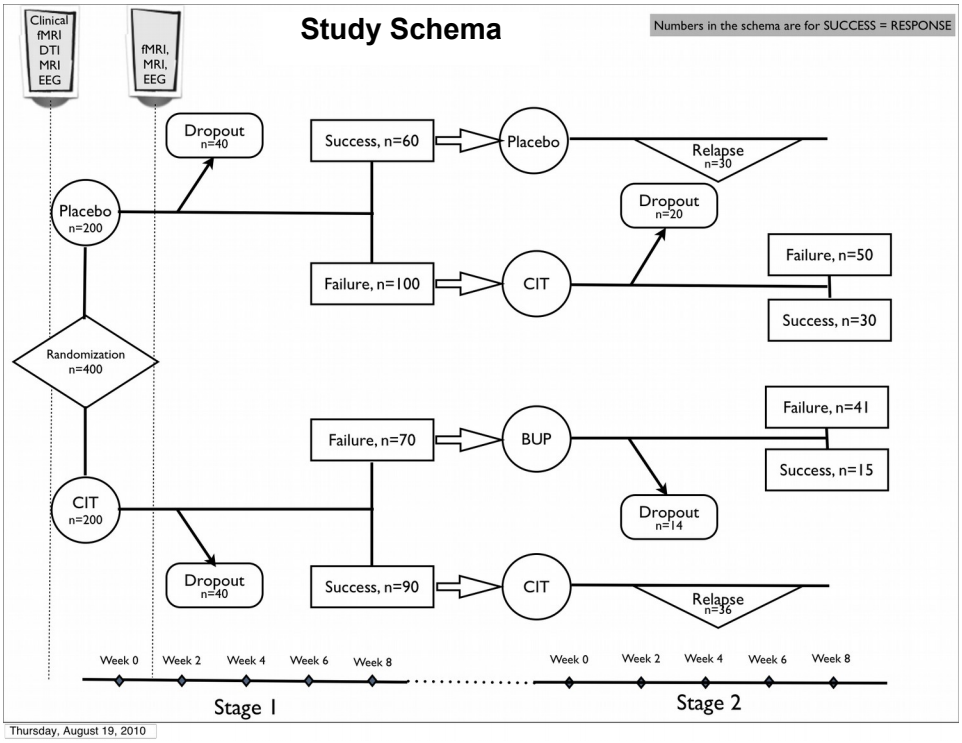
5.14 DATA ANALYSIS PLAN

5.14.1. Study design and Aims

Aim 1: To identify baseline clinical, neuroimaging, neurophysiological, and behavioral, moderators of differential treatment outcome (mean symptom change and tolerability) for CIT versus PBO for the treatment of MDD.

Aim 2: To identify early phase (week 1) neuroimaging, neurophysiological, and behavioral mediators of differential treatment outcomes (symptom change, tolerability) to CIT versus PBO.

Aim 3: To compare the 8-week outcomes of CIT vs. PBO based on rates of symptom remission, defined as a score of 7 or less on the 17-item Hamilton Rating Scale for Depression ($HRSD_{17}$).[54]



Aim 4: To develop methods to examine the complex interactions of clinical, neuroimaging, neurophysiological, and behavioral moderators and mediators of differential treatment outcomes (mean symptom change, tolerability) for CIT and PBO.

Aim 5: To develop a depression differential treatment response index (DTRI) that integrates moderators across clinical and biological variables.

Aim 6: To develop a composite scale of treatment acceptability (which incorporates depressive symptom outcomes, treatment tolerability, and the patient's functional status) and compare 8-week outcomes for CIT and PBO.

5.14.2 Outcomes

5.14.2.1 Depression symptoms

(a) Primary: course of symptoms during treatment (measured by the regression slope over time)
(b) Secondary: response status at treatment end
(c) Secondary: remission status at treatment end

5.14.2.2 Side effects

(a) Frequency, Intensity, and Burden of Side Effects Rating (FIBSER)
(b) SAFTEE

5.14.3 Candidate moderators and mediators

5.14.3.1 Moderators

The following measures assessed at baseline will be considered as potential moderators of outcome
(a) Clinical phenotypes
(b) Electrophysiological measures (EEG)
(c) Behavioral measures [psychomotor slowing, cognitive control, working memory, and reward responsiveness]
(d) Structural brain measures (MRI and DTI)
(e) Functional brain measures, activation during task and resting state connectivity (fMRI)

5.14.3.2 Mediators

The following changes from baseline to week one will be considered as potential mediators of outcome

(a) Electrophysiological measures (EEG)
(b) Structural brain measures (MRI)
(c) Functional brain measures, activation during task and resting state connectivity (fMRI)

5.14.4 Data analytic considerations

5.14.4.1 Definitions

For the purposes of this application, the following definitions are adopted

(a) *Moderator*: baseline characteristics, such that the association between the characteristic and the outcome is different between the two treatments. Specifically, we investigate looking for moderators that have strong association with the outcome among CIT-treated patients and no or a weak association with the outcome among placebo treated patients.

(b) *Mediator*: a variable measured post-randomization, but before the outcome (i.e. intermediate variable), such that the association between this intermediate variable and the outcome is different between the two treatments. Specifically, investigate for mediators measured 1-2 weeks post-randomization, that have strong association with the outcome among CIT-treated patients and no or a weak association with the outcome among placebo treated patients.

5.14.4.2 Site effects

Procedures are in place to ensure precise adherence to the study protocol (a comprehensive training plan and utilization of web-based trial management software) and comparable symptoms rating criteria (reliability training plan) at all study sites, to minimize the heterogeneity between sites. In addition, plan will be put in place for real-time monitoring for drift in ratings and measuring biosignatures at all sites and if necessary, re-training of raters and technicians who show deviations will be done. Prior to all analyses we will examine all measures for evidence for systematic differences between sites. All measures at all sites will be described in terms of means, medians, standard deviations, minimums and maximums, and 95% CI for the means and proportions. The variation between sites with respect to the associations between treatment outcome and all biosignatures will be documented and the selection of predictors will take into account this variability. Testing of the study hypotheses will account for the potential site effect by including a random site effect.

5.14.4.3 Dropout

We expect that most dropout will occur during the initial washout period when some participants will experience improvement with supportive clinical management prior to beginning medication treatment. Those who are randomized and enter the treatment are expected to be sufficiently motivated and less likely to drop out. Nevertheless, we have established procedures for tracking patients who drop out and making a final assessment on them including reasons for dropout, including payment for completing the assessment. We expect about 20% dropout. Data from patients with at least 2 observations will be utilized, since the primary outcome is course of symptoms during treatment, modeled with mixed effects models for longitudinal data, see model (1) below. The collected reasons for dropout will be used for assessing the acceptability of treatment.

5.14.4.3 Placebo effects

A problem in all antidepressant studies is the fact that some subjects improve due to nonspecific aspects of the treatment, so called “placebo effects”, rather than the due to the specific effect of the treatment, which would be the chemical component in the drug. In randomized clinical trials such subjects are equally present in the drug and in the placebo arm. The presence of such individuals in the drug arm makes the identification of moderators and mediators more difficult, since while the outcome in the placebo group consists only of placebo effects, the outcomes in the drug group is not the “pure” result of the drug, but rather a combination of drug and placebo effects.

Our group has expertise in dealing with this problem. We have developed a method for estimating, on an individual basis, the outcome that is due to placebo effects when the subject is treated with an active drug.[207] In addition to the standard analysis, where in model (1) below we will model the observed HAMD scores during treatment, we will also apply the above methodology to estimate for each individual in the drug group, her/his specific drug effect and use that as outcome in (1).

5.14.5 Analytic strategy

5.14.5.1 Assessment of individual moderators/mediators

With the primary outcome the following mixed effects model will be used to identify treatment moderators:

$$Y_{it} = \beta_0 + \beta_1 trt_i + \beta_2 t + \beta_{12} t * trt_i + \beta_3 X_i + \beta_{13} trt_i * X_i + \beta_{23} t * X_i + \beta_4 t * trt_i * X_i + error_{it},$$

where Y_{it} indicates the depression severity measure for subjects i at time t, trt_i is an indicator for the treatment for subject i (for stage 1 data it would be an indicator for CIT (vs. placebo)) and X_i is the value of the variable,

candidate for a moderator of treatment outcome for subject *i*. The term $error_{it}$ consists of random effects for subjects' individual slopes and intercept and a random error term with mean zero. A significant β_4 term would identify *X* as a moderator.

With the secondary outcomes (remission and response), the following models will employed:

$$logit[P(Y=1)] = \beta_0 + \beta_1 trt_i + \beta_2 X_i + \beta_{12} trt_i * X_i \ ,$$

with the same notation as above. A significant β_{12} coefficient will indicate that *X* is a moderator.

Identifying mediators will employ similar models.

Based on those analyses we will rank the moderators according to some criteria. One criterion is the effect size of the difference between drug and placebo with respect to change in the outcome for one standard deviation change in the moderator X. An alternative is to construct receiver operating characteristics (ROC) curve for, say, outcome=response separately for drug (CIT) and for placebo and take the difference between the area under the curve of the two ROCs. Ranking will inform the importance of each of the variable alone as a biomarker.

5.14.5.2 Combining biomarkers

The second step is to combine baseline variables to create a composite measure that has best properties of a moderator. (Similar strategy will be employed for selecting intermediate variables as predictors of outcome after treatment begins.) This analytic step will involve multivariate methods and variables selections approaches. Because of the large number of potential predictors (p, in the order of possibly thousands) relative to the number of cases (n, here total 400), the major challenge in the analysis of the data to be collected is to build a statistical model with extremely high dimensional (hundreds of thousands) predictor space. For a continuous outcome, the natural model to consider is the linear regression model,

$$y_i = \alpha + \sum_{j=1}^{N} \beta_j x_{ij} + \varepsilon_{i,} \ i = 1,\ldots,n \text{, or equivalently, in matrix form } Y = X\beta + \varepsilon \qquad (1)$$

where y_i is the response from the *i*th subject, $x_{i1},\ldots,x_{iN}$ are the predictors for the same subject, and ε_i is an error term, with $Y = (y_i,\ldots,y_n)^T$, $\varepsilon = (\varepsilon_1,\ldots\varepsilon_n)^T$, and the matrix *X* contains the predictors and a column of ones for the intercept. In the data we propose to collect, however, the number of potential predictors *N* is much larger than the number of observations *n*, and thus classical modeling approaches are not appropriate– this is known as the "small n large p" problem. In order to provide for a valid and interpretable model, some sort of dimension reduction is necessary.

Popular methods for variable selection (e.g., forward selection, backwards deletion, and stepwise methods) are impractical when there are many potential predictors. Similarly, information criteria (e.g., Akaike's information criterion (AIC) and the Bayesian information criterion (BIC)) are not equipped to deal with huge numbers of predictors such as what we will encounter. Here we outline various general strategies that can be used to construct models for treatment outcomes using the various biosignatures (fMRI data, EEG data, cortical thickness, etc.) to be gathered.

(a) Principal component regression and related methods

Dimension reduction has been an important objective in general multivariate statistical analysis. One of the most common methods for this is principal component analysis (PCA) which consists of transforming the high-dimensional data to a new set of coordinate axes, in which most of the variation in the data is expressed in the first few components.

Performing regression analysis in this transformed space, known as principal component regression, is a popular and often effective method for reducing dimensionality in the general model (1). Alternatively, rather than selecting components as to how much of the variation in the predictors is accounted for (as in PCR), components may be selected according to how well they predict the response variable. The method of partial least squares (PLS)[208-210] is one approach to this, choosing orthogonal components iteratively to maximize the covariance between the components and the response.

Other related approaches to decreasing the dimensionality by transforming the predictor space include "sliced inverse regression" (SIR)[211], Sliced Average Variance Estimation (SAVE)[212], and Minimum Average Variance Estimation (MAVE)[213]. Yet another alternative is independent component analysis (ICA),[214] which chooses components of the predictor data that maximize an index of mutual statistical independence. Like principal components, independent components are not necessarily related to the outcome, but unlike the

former they are usually expected to represent distinct physically meaningful sources of variability. Thus, whereas restriction to leading principal components is merely a means to the end of dimension reduction, regression on independent components may be useful for assessing the impact of physically distinct sources of variation on the outcome.

(b) Machine learning algorithms

Recently there has been much activity in the area of machine learning, i.e., developing algorithms capable of recognizing complex patterns in data. A typical example of such algorithms involves classifying high dimensional observations into two classes. Popular methods include various forms of discriminant analysis and related classification methods[215], neural networks[216], and support vector machines (SVM)[217], overviewed by Hastie et al.[218].

Such approaches have been taken in a variety of application areas relevant to this proposal, including in using imaging data (typically fMRI) to discriminate clinically relevant groups such as a discriminative analysis procedure based on nonlinear SVMs[219], principal component analysis followed by Fisher's linear discriminant (FLD) classification[220], comparisons between SVM and FLD for classifying activation patterns[221], use of ROI data in an fMRI experiment to predict a subject's perceptual decision[222], a combination of SVM and maximum uncertainty linear discriminant analysis (MLDA) to predict mental state using data from fMRI experiments[223], and SVM-based prediction of depression diagnosis using data from an fMRI task[224].

Inasmuch as the aim is to predict new observations based on available biosignatures, such methods are promising. They are less satisfactory, however, when interest lies more in understanding the relationship between clinical outcomes and the various predictors.

(c) Tree-based methods

Another alternative modeling strategy involves partitioning the predictor space into many rectangles, then fitting a simple model to the data that fall in each. This algorithm works recursively, considering all variables in turn and selecting split points that provide the best fit for the response variables. This approach is termed "classification and regression trees" (CART)[225], applicable both when the response variable is binary (e.g., remitted/not remitted) or continuous (e.g., measures of depression severity). Related to CART is multivariate adaptive regression splints (MARS).[226] A number of newer innovations have been made, including aggregating trees across many bootstrap samples of the data ("bagging")[227] and boosting algorithms.[228]

(d) Penalization

An alternative to variable selection methods mentioned earlier that has proven useful in nonparametric regression and functional data analysis is the idea of reducing dimensionality at the same time as fitting the model by introducing a penalty for model complexity into the optimization. This has been particularly useful when the dimensionality is high, as it simultaneously chooses a model of reduced dimension and fits it. If *N* is small relative to *n* then the model (1) may be fit simply by minimizing the usual least squares criterion, i.e., choosing β to minimize $(y - X\beta)^T (y - X\beta)$. However, in our application, *N* will be (much) larger than *n* and therefore, as mentioned earlier, this simple approach is not appropriate since the optimization problem is ill-posed. To enforce some level of sparsity to the resulting model for the data, we might consider applying an L^1 (sum of absolute value) penalty to the coefficient vector, i.e., minimizing instead

$$(y - X\beta)^T (y - X\beta) + \lambda \sum_{j=1}^{N} |\beta_j| \qquad (2)$$

over all choices of β. Minimizing (2) will accomplish both variable selection and estimation simultaneously, as many of the β_j values will be set to zero, effectively deselecting the corresponding predictors from the model. This is the well-known "Least Absolute Shrinkage and Selection Operator" (LASSO) method,[229] which may be fit using the efficient LARS algorithm.[230] There are a number of related alternatives to the L^1 penalization method, including the smoothly clipped absolute deviation (SCAD) penalty,[231] the adaptive lasso,[232] and the Dantzig selector.[233]

Applying methods of this type necessitate the choice of a tuning parameter λ as in (2). This controls the amount of penalization that will be applied to model size: a large value of λ will apply a relatively severe penalty, tending to result in models with relatively few predictors; a small value of λ will naturally result in more predictors. There are a number of data-adaptive methods for selecting an appropriate value of λ, including cross-validation (CV), generalized CV and other related techniques,[234] which we will adapt to our situation here.

(e) Wavelet analysis

One useful approach in the analysis of data consisting of spectra, curves, or images is to project data onto a fixed set of basis functions and choose some subset of these to represent the data. Historically popular basis functions include Fourier bases and orthogonal polynomials. In such an analysis it is typical to represent the function by truncating the series expansion at some point. This not only reduces the dimensionality of the data but also enforces some measure of smoothness on it.

More recently, the field of wavelet analysis has been developed and applied in a wide variety of fields, including the analysis of time series or spectral/imaging data. For the purposes of the grant we favor such an approach because wavelets allow for a sparse representation of such data and because it is well suited for representing such data with a mixture of large- and small-scale features. Related to this is their ability to efficiently represent jumps, spikes, and peaks in functions and images, and thus they have become an important tool in nonparametric regression and other related statistical techniques.[235, 236]

A family of wavelets is created from a single "mother wavelet" function ψ by dilating and translating: $\psi_{jk}(t) = 2^{j/2}\psi(2^j t - k)$, where *j* is the dilation index (which controls the scale of the resulting wavelet function) and *k* is the translation index (which specifies the location). Of particular interest in statistical applications are orthonormal and compactly supported wavelet families.[237] Orthonormality is important in many statistical applications since an orthonormal wavelet transform of independent data preserves independence. Even when the observed data are not independent, wavelet transforms have a natural "whitening" property[236] and thus have become popular in brain imaging applications (among many others).

Signals and images are converted to the "wavelet domain" by means of the discrete wavelet transform (DWT), which consists of iterative application of discrete filtering operations. However, simply transforming the data to the wavelet domain does not provide any dimension reduction. Reduction can be accomplished by exploiting the sparse wavelet representation of data (i.e., most coefficients are essentially zero), in that data are well approximated using only a relatively small number of wavelet functions. When using wavelets in the nonparametric regression model, dimension reduction is typically achieved by using some form of wavelet thresholding, e.g., choosing some relatively small number of wavelet functions the wavelet functions according to the magnitudes of their associated coefficients.

Wavelets, originally developed in one dimension, now enjoy widespread application in image and video processing, utilizing their extensions to two and three dimensions. In two dimensions, taking tensor products of the mother wavelet and the scaling function result in three different two-dimensional wavelet functions, roughly representing "vertical", "horizontal", and "diagonal" detail. The discrete wavelet transform for two dimensions is accomplished by recursive application of the usual one-dimensional wavelet filters along both the rows and the columns of the array representing the image. Similarly, a three-dimensional wavelet transform can be conducted by applying the filtering operations along the third dimension as well.

(f) Functional data analysis

Functional data analysis (FDA)[240] is a relatively new direction in which statistical methodology may be generalized, by extending statistical models for vectors to data which can be represented as curves, functions, or images. This is relevant to this project since many of the biosignatures that are being collected (e.g., imaging data) can be regarded as functional data, and thus some of these techniques are needed to deal with them adequately.

Since the early 1990s, many authors have developed new versions of many classical statistical techniques, such as analysis of variance (ANOVA) and principal component analysis, to accommodate functional observations. Although most of the FDA literature has dealt with functions of one variable, others have employed the FDA paradigm in the study of two- or three-dimensional images,[241] which may be regarded as functions of more than one variable.

The major breakthrough accomplished by FDA is the development of fairly general approaches to handling very high-dimensional data sets even with small or modest sample sizes. These methods all exploit the particular structure of the functional data — in the case of brain imaging data, for instance, though there may be hundreds of thousands of measurements made, there is strong local correlation throughout, due in part to similarity of physiological characteristics of brain regions and in part to correlation of noise possibly due to partial volume and other effects. And thus in FDA applications the "effective dimensionality" is much smaller than the apparent dimensionality and so all FDA methods necessarily involve some form of dimension reduction.

Particularly relevant to this research is functional linear regression, specifically situations in which the response variable (e.g., remission to treatment) is scalar and the response variables include functional data, the fastest-growing area of FDA in the last decade.[240] Recent developments include functional logistic regression,[242] further generalization of the response variables,[243] and functional adaptive model estimation,[244] which generalizes both functional GLM and generalized projection pursuit.[245]

For a set of one-dimensional functional predictors $X_1(t),\mathrm{K},X_n(t)$ (e.g., the spectral content of an EEG signal) with corresponding continuous responses y_1,K,y_n (e.g., post-treatment HAM-D scores), the simplest functional linear regression model may be expressed $y_i = \alpha + \int X_i(t)\omega(t)dt + \varepsilon_i,\ i = 1,\mathrm{K},n$. The interpretation of the coefficient function ω is that areas in which $\omega(t)$ is positive (negative) indicates that larger values of the signal in that area tend to correspond with larger (smaller) responses. Similarly, if $\omega(t)$ is near zero in some area, the value of the signals in that area have little effect on the predictor.

The main challenge is in estimating the coefficient function ω. In practice, the n signals are sampled at a finite number (*N*, say) of points — the problem thus resembles the general model given in (1), i.e., fitting *n* observations with *N* possible predictors. Taking an FDA approach involves reducing the dimensionality of the predictors while exploiting the structure of the functional data. This typically involves some sort of penalization, e.g., penalizing the roughness of the estimated ω. Our approach will involve first transforming the functional data (spectra, images) into the wavelet domain, then considering the various approaches outlined in this section, providing for appropriate FDA-type treatment of the data.

(g) Variable screening

In practice, the penalization and other approaches described here work well even for rather high dimensional predictors. However, when considering "ultra-high-dimensional" predictors as in the data we propose to collect, there may be some advantage gained by first examining all potential predictors and screening out those that are judged not likely to help in predicting the outcome.

As an example, Fan and Lv (2008)[246] propose a two-step procedure called "sure independence screening" (SIS). In the first step one pre-screens the predictors, retaining a manageable number of predictors most highly correlated with the outcome. In the second step, a criterion such as (2) is optimized to fit the "final" model. Work currently in process by the investigators has found that such a method is indeed more effective than not prescreening predictors when the dimensionality of the predictor space is very high. Furthermore, there is some preliminary evidence that it may be even more advantageous to screen on the basis of the amount of variability in the images accounted for by each coefficient, rather than the correlation between the coefficients and the outcome variable.

We note that such screening procedures are meant for general predictors in which there may not be any particular relationship among the predictors that could be exploited. In our situation, however, much of the data we will gather for each subject are functional (e.g., the imaging data), and therefore such a method may be fine-tuned to take advantage of the natural structure of the data. As mentioned before, all functional predictors will be transformed into the wavelet domain and the predictors thus consist of the wavelet coefficients. We note that these coefficients are naturally grouped according to scale — for a one-dimensional (discretely sampled) function, half the coefficients correspond to the finest scale features, one-fourth to the next finest scale, one-eighth to the 3rd level of detail, and so on. If we are willing to assume that the true coefficient function predominantly consists of relatively large-scale features, then we can improve performance by adjusting the screening threshold for each level, i.e., "raising the bar" for the higher level coefficients. The imbalance of the scale coefficients is greater for images: for a two-dimensional image, ¾ of the coefficients are at the highest level of detail, 3/16 at the next highest level, etc.

(i) Generalized linear models

Thus far this section has dealt only with situations in which the outcome variable is continuous, but we are also interested in situations in which the outcome is binary, e.g., remitted or not remitted. This falls under the classification of general linear models (GLM[247]), which extend linear models to any response variable *Y* whose density can be written as $p(y;\theta,\phi) = \exp[(y\theta - b(\theta))/a(\phi)] + c(y,\phi)]$. The response is related to the predictor through the model $g(\mu) = \alpha + \beta^T x_i$ for some function *g*, where _ $\mu = E(Y;\theta,\phi)$. The function *g* is termed the link function since it connects the linear predictor with the expected response.

The logistic regression model,[248] one of the most important of the GLMs, has found broad application in biomedical and other settings. This model predicts the probability of a binary outcome, such as remission from

MDD, as a function of covariates through the equation $\log(p_i/(1-p_i)) = \alpha + \beta^T x_i$, where p_i is the event's probability. Typically, in medical settings, the data for each subject consist of *k* clinical covariates along with the response Y_i, which equals 1 if the event of interest occurs and 0 otherwise.

Extending functional regression models from the linear to the generalized linear case entails several methodological hurdles. For instance, the penalized least squares criterion minimized in functional PCR/PLS must, in the GLM case, be minimized at each iteration of the iteratively reweighted least squares algorithm used to fit GLMs. Moreover, the criteria for selecting the smoothing parameters must be modified in the GLM case.

For fitting functional GLMs when the predictor space is very high dimensional, as for the case with continuous outcomes, we will apply an L^1-type penalty to the coefficients in order to ensure sparsity of the solution. To accomplish this, the regularization path algorithm[249], which provides an efficient means of fitting L^1 penalties will be adapted to our situation.

(j) Comparison and validation of models

In order to compare the various modeling approaches we will apply them each to the gathered data. First, we will randomly divide the data for each treatment group into "training" (consisting of 2/3rd of the subjects for each treatment group) and "validation" subsets (the remaining 1/3rd of the subjects for each group). The model will be fit, including the selection of tuning parameters, to the training datasets. Then the various methods may be compared as to how well they predict outcomes in the corresponding validation subsets. The utility of the various fitted models will be measured according to two different criteria depending on the nature of the outcome variable in each situation.

Prediction error. In the regression model with a continuous (or near-continuous) outcome (e.g., measure of depression severity, slope of the trajectory of symptoms change during treatment), the prediction error is defined as the error made when predicting a "new" observation (e.g., a patient outcome) using his/her biosignature data. Other factors being equal, we would favor a method that results in smaller prediction error.

Classification error**.** In applications with binary outcomes (remission/nonremission), performance of the methods will be evaluated based on the sensitivity (proportion of positives classified as positive) and specificity (proportion of negatives classified as negative). The ROC curves will be constructed and different models wil be compared with respect to the difference between area under the curve for drug and placebo. This will allow for direct comparisons of all the various approaches described here.

5.14.5.3 Developing DTRI and improvement/side effects matrix

Based on the univariate examination of all candidate moderators of treatment outcome (see 5.14.5.1) and statistical modeling for identifying a combination of the candidates covariates that best moderate treatment outcome (see 5.14.5.2), indices associated with depression symptoms improvement and side effects associated with CIT and placebo will be developed.

The depression treatment response indices (DTRI), computed at baseline, will be associated with levels of improvement expected with CIT treatment (DTRI-CIT). Improvement might be measured in several different ways (e.g., absolute improvement from baseline, or improvement on the drug, relative to improvement on placebo for the same level of the index).

The side effect indices Index-SE, also computed pre-treatment, will be associated with the severity of side effects expected with the specific treatment, with different indices for CIT and BUP. Severity of the side effects can be measured by FIBSER, SAFTEE or a combination of them.

Individuals with the same level of the index will be expected to have equal improvement (or severity of side effects). The DTRI index can be matched to different outcomes: for example, probability of response or remission, or % or absolute improvement from baseline. In practice, the indices can be constructed in 2 ways. In the first approach points can be assigned for different ranges of the variables used in the selected model and their combinations; the points then can be summed and the sum will constitute the index. The pros of this approach is the simplicity of computation, allowing doctors and patients to understand how different factors affect patients' probability to respond to a given treatment. The Farmingham Cardiac Risk score is constructed in this way. We call this a *points-counting* method. The second approach is directly based on the best model and, for given values of the moderators in this model, the linear combination of main effects and interactions is computed (on the background by a computer). We call this approach a *black-box* method since clinicians and patients do not directly see how the index is constructed, rather they are provided only with the final index. Pro of this method is accuracy of the index (for example, continuous moderators need not be categorized in order

to determine the number of points added to the index, as in the first approach); con is the need to have a software in clinicians' offices that would compute the index based on the values of the moderators; of course, an index calculator can be made available on line.

Prior to treatment, patients can be charted on a graph as the one depicted on Figure 4 below. Subjects falling in the lower left corner would be good candidates for treatment with CIT. Guidelines for treatment can be established based on the chart.

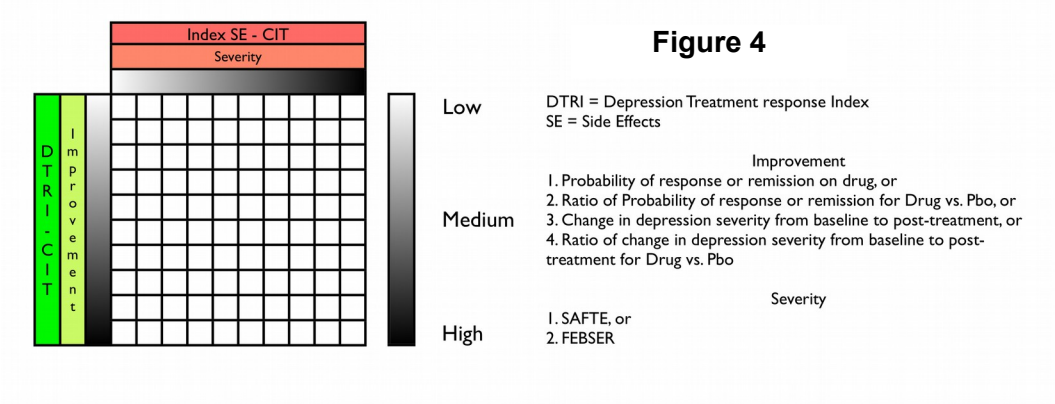
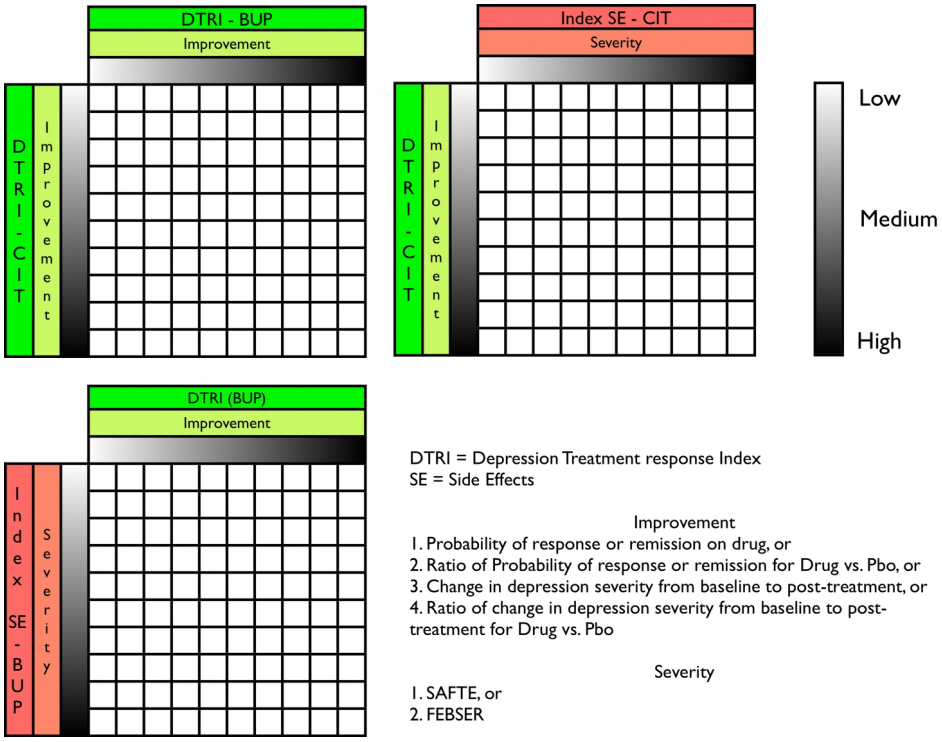


Figure 4

Although the design of the proposed study does not make possible the similar development of BUP indices, it does allow to begin the process of identifying baseline factors that are specifically related to outcome from treatment with BUP. Comparing the relationship between outcome from BUP treatment and baseline predictors for subjects unsuccessfully treated with CIT in Stage 1 vs. the relationship between the same baseline predictor and outcome for placebo treated subjects in Stage 1, would suggest potential moderators of BUP treatment response. While the two groups are not exactly comparable, this comparison would give an indication for the baseline characteristics that are related to outcome with BUP treatment among failures to treatment with CIT. It will also yield results that can be useful leads in future studies, which is a major goal of this project which aims to generate candidate moderator and mediator variables rather than to definitely test *a priori* hypotheses. Helping this comparison is the fact that the BUP treated group of patients failing CIT treatment would contain fewer placebo responders, thus increasing the signal in the comparison against placebo. Ultimately, a DTRI and a side effects index for BUP will be used in combination with the similar indices for CIT. Prior to treatment patients can be charted on a graph as the one depicted on Figure 5 below. Using personal judgment or guidelines, the clinician can make a decision what treatment to offer to a depressed patient. For example, if a patient falls in the top left corner on the (DTRI-CIT, DTRI-BUP) graph, it means that s/he has is expected to have low improvement on both CIT and BUP, and the clinician might chose an antidepressant with a different mechanism of action. If a patient falls on the right diagonal (\) of the (DTRI-CIT, DTRI-BUP) graph, this would mean that her/his improvement is expected to be the same on CIT and BUP and in this case side effects considerations from (DTRI-CIT, Index-SE CIT) and (Index-SE BUP, DTRI-BUP) graphs can help decide the treatment.

Figure 5.



Ancillary Analyses: During the course of a trial a number of ancillary analyses will be identified which are not part of the study.

5.15. Resource Sharing Plan:

We will send all genetic, transcriptomic, and proteomic samples to the repository for processing and storage. It will also house the MRI, and EEG data. The repository will be responsible for distributing data, to other qualified investigators as directed by the PIs. We will send all genetic and proteomic samples to the NIMH repository at Rutgers University for processing and storage, which will also serve as repository for the clinical, MRI, and EEG data. The NIMH repository at Rutgers University was established in 1999 as a national resource of clinical data and biomaterials that are collected from individuals with major neuropsychiatric disorders. It has housed such data and biomaterials (cell lines and DNA samples) to be available to qualified investigators who study the genetics of psychiatric disorders. For example, data and biomaterials were collected in six projects that participated in the National Institute of Mental Health (NIMH) Genetics of Recurrent Early-Onset Depression (GenRED) project, from 1999-2003, which included among its Principal Investigators and Co-Investigators one of our PI's, Myrna Weissman, and the investigator on this project who will chair our genetics core, Dr. James Knowles. The repository has the capacity for the culture of immortalized cell lines to provide DNA specimens, the capacity for secure de-identified long-term storage of other biologic materials, and the secure storage or digital data related to NIMH sponsored projects, and provision of samples to qualified investigators with NIMH approval. In addition, we will also submit all DNA sequence data from the RNA-Seq experiments to the Sequence Read Archive (SRA), or any other public database the NIMH requires.

REFERENCES:

1. Murphy, S.A., K.G. Lynch, D. Oslin, J.R. McKay, and T. TenHave, *Developing adaptive treatment strategies in substance abuse research.* Drug Alcohol Depend, 2007. **88 Suppl 2**: p. S24-30.
2. Kraemer, H.C., G.T. Wilson, C.G. Fairburn, and W.S. Agras, *Mediators and moderators of treatment effects in randomized clinical trials.* Arch Gen Psychiatry, 2002. **59**(10): p. 877-83.
3. Fava, M., A.J. Rush, M.H. Trivedi, A.A. Nierenberg, M.E. Thase, H.A. Sackeim, F.M. Quitkin, S. Wisniewski, P.W. Lavori, J.F. Rosenbaum, and D.J. Kupfer, *Background and rationale for the sequenced treatment alternatives to relieve depression (STAR*D) study.* Psychiatric Clinics of North America, 2003. **26**(2): p. 457-494.
4. Rush, A.J., M. Fava, S.R. Wisniewski, P.W. Lavori, M.H. Trivedi, H.A. Sackeim, M.E. Thase, A.A. Nierenberg, F.M. Quitkin, and T.M. Kashner, *Sequenced treatment alternatives to relieve depression (STAR*D): rationale and design.* Controlled Clinical Trials, 2004. **25**(1): p. 119-142.
5. Trivedi, M.H., A.J. Rush, S.R. Wisniewski, A.A. Nierenberg, D. Warden, L. Ritz, G. Norquist, R.H. Howland, B. Lebowitz, P.J. McGrath, K. Shores-Wilson, M.M. Biggs, G.K. Balasubramani, M. Fava, and 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**(1): p. 28-40.
6. Carlson, P.J., J.B. Singh, C.A. Zarate, Jr., W.C. Drevets, and H.K. Manji, *Neural circuitry and neuroplasticity in mood disorders: insights for novel therapeutic targets.* NeuroRx, 2006. **3**(1): p. 22-41.
7. Krishnan, V. and E.J. Nestler, *The molecular neurobiology of depression.* Nature, 2008. **455**(7215): p. 894-902.
8. Nemeroff, C.B. and W.W. Vale, *The neurobiology of depression: inroads to treatment and new drug discovery.* J Clin Psychiatry, 2005. **66 Suppl 7**: p. 5-13.
9. Shelton, R.C., *The molecular neurobiology of depression.* Psychiatr Clin North Am, 2007. **30**(1): p. 1-11.
10. Papakostas, G.I. and M. Fava, *Predictors, moderators, and mediators (correlates) of treatment outcome in major depressive disorder.* Dialogues Clin Neurosci, 2008. **10**(4): p. 439-51.
11. Olfson, M. and S.C. Marcus, *National patterns in antidepressant medication treatment.* Arch Gen Psychiatry, 2009. **66**(8): p. 848-56.
12. Owens, M.J., D.L. Knight, and C.B. Nemeroff, *Second-generation SSRIs: human monoamine transporter binding profile of escitalopram and R-fluoxetine.* Biol Psychiatry, 2001. **50**(5): p. 345-50.
13. Bolden-Watson, C. and E. Richelson, *Blockade by newly-developed antidepressants of biogenic amine uptake into rat brain synaptosomes.* Life Sci, 1993. **52**(12): p. 1023-9.
14. Koch, S., K.W. Perry, D.L. Nelson, R.G. Conway, P.G. Threlkeld, and F.P. Bymaster, *R-fluoxetine increases extracellular DA, NE, as well as 5-HT in rat prefrontal cortex and hypothalamus: an in vivo microdialysis and receptor binding study.* Neuropsychopharmacology, 2002. **27**(6): p. 949-59.
15. Ascher, J.A., J.O. Cole, J.N. Colin, J.P. Feighner, R.M. Ferris, H.C. Fibiger, R.N. Golden, P. Martin, W.Z. Potter, E. Richelson, and et al., *Bupropion: a review of its mechanism of antidepressant activity.* J Clin Psychiatry, 1995. **56**(9): p. 395-401.
16. Meyer, J.H., V.S. Goulding, A.A. Wilson, D. Hussey, B.K. Christensen, and S. Houle, *Bupropion occupancy of the dopamine transporter is low during clinical treatment.* Psychopharmacology (Berl), 2002. **163**(1): p. 102-5.
17. Learned-Coughlin, S.M., M. Bergstrom, I. Savitcheva, J. Ascher, V.D. Schmith, and B. Langstrom, *In vivo activity of bupropion at the human dopamine transporter as measured by positron emission tomography.* Biol Psychiatry, 2003. **54**(8): p. 800-5.
18. Szabo, Z., M. Argyelan, B. Kanyo, L. Pavics, and Z. Janka, *[Change of dopamine transporter activity (DAT) during the action of bupropion (in depression)].* Neuropsychopharmacol Hung, 2004. **6**(2): p. 79-81.
19. Dong, J. and P. Blier, *Modification of norepinephrine and serotonin, but not dopamine, neuron firing by sustained bupropion treatment.* Psychopharmacology (Berl), 2001. **155**(1): p. 52-7.
20. Li, S.X., K.W. Perry, and D.T. Wong, *Influence of fluoxetine on the ability of bupropion to modulate extracellular dopamine and norepinephrine concentrations in three mesocorticolimbic areas of rats.* Neuropharmacology, 2002. **42**(2): p. 181-90.

21. Nomikos, G.G., G. Damsma, D. Wenkstern, and H.C. Fibiger, *Acute effects of bupropion on extracellular dopamine concentrations in rat striatum and nucleus accumbens studied by in vivo microdialysis.* Neuropsychopharmacology, 1989. **2**(4): p. 273-9.
22. Bondarev, M.L., T.S. Bondareva, R. Young, and R.A. Glennon, *Behavioral and biochemical investigations of bupropion metabolites.* Eur J Pharmacol, 2003. **474**(1): p. 85-93.
23. Rush, A.J., M.H. Trivedi, S.R. Wisniewski, A.A. Nierenberg, J.W. Stewart, D. Warden, G. Niederehe, M.E. Thase, P.W. Lavori, B.D. Lebowitz, P.J. McGrath, J.F. Rosenbaum, H.A. Sackeim, D.J. Kupfer, J. Luther, and M. Fava, *Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D report.* Am J Psychiatry, 2006. **163**(11): p. 1905-17.
24. Anguiano, A., J.R. Nevins, and A. Potti, *Toward the individualization of lung cancer therapy.* Cancer, 2008. **113**(7 Suppl): p. 1760-7.
25. Dowsett, M. and A.K. Dunbier, *Emerging biomarkers and new understanding of traditional markers in personalized therapy for breast cancer.* Clin Cancer Res, 2008. **14**(24): p. 8019-26.
26. Lee, A.Y., A.K. Raya, S.M. Kymes, A. Shiels, and M.A. Brantley, Jr., *Pharmacogenetics of complement factor H (Y402H) and treatment of exudative age-related macular degeneration with ranibizumab.* Br J Ophthalmol, 2009. **93**(5): p. 610-3.
27. Lima, J.J., K.V. Blake, K.G. Tantisira, and S.T. Weiss, *Pharmacogenetics of asthma.* Curr Opin Pulm Med, 2009. **15**(1): p. 57-62.
28. Ouzounian, M., D.S. Lee, A.O. Gramolini, A. Emili, M. Fukuoka, and P.P. Liu, *Predict, prevent and personalize: Genomic and proteomic approaches to cardiovascular medicine.* Can J Cardiol, 2007. **23 Suppl A**: p. 28A-33A.
29. Vosslamber, S., L.G. van Baarsen, and C.L. Verweij, *Pharmacogenomics of IFN-beta in multiple sclerosis: towards a personalized medicine approach.* Pharmacogenomics, 2009. **10**(1): p. 97-108.
30. Holsboer, F. and N. Barden, *Antidepressants and hypothalamic-pituitary-adrenocortical regulation.* Endocr Rev, 1996. **17**(2): p. 187-205.
31. Kraemer, H.C., E. Stice, A. Kazdin, D. Offord, and D. Kupfer, *How do risk factors work together? Mediators, moderators, and independent, overlapping, and proxy risk factors.* Am J Psychiatry, 2001. **158**(6): p. 848-56.
32. Nierenberg, A.A., *Predictors of response to antidepressants general principles and clinical implications.* Psychiatr Clin North Am, 2003. **26**(2): p. 345-52, viii.
33. Rush, A.J. and R.F. Prien, *From scientific knowledge to the clinical practice of psychopharmacology: can the gap be bridged?* Psychopharmacol Bull, 1995. **31**(1): p. 7-20.
34. Rush, A.J., S.R. Wisniewski, D. Warden, J.F. Luther, L.L. Davis, M. Fava, A.A. Nierenberg, and M.H. Trivedi, *Selecting among second-step antidepressant medication monotherapies: predictive value of clinical, demographic, or first-step treatment features.* Arch Gen Psychiatry, 2008. **65**(8): p. 870-80.
35. Fava, M., J.E. Alpert, C.N. Carmin, S.R. Wisniewski, M.H. Trivedi, M.M. Biggs, K. Shores-Wilson, D. Morgan, T. Schwartz, G.K. Balasubramani, and A.J. Rush, *Clinical correlates and symptom patterns of anxious depression among patients with major depressive disorder in STAR*D.* Psychol.Med., 2004. **34**(7): p. 1299-1308.
36. Fava, M., M.A. Rankin, E.C. Wright, J.E. Alpert, A.A. Nierenberg, J. Pava, and J.F. Rosenbaum, *Anxiety disorders in major depression.* Compr Psychiatry, 2000. **41**(2): p. 97-102.
37. Fava, M., *Psychopharmacologic treatment of pathologic aggression.* Psychiatr Clin North Am, 1997. **20**(2): p. 427-51.
38. Fisar, Z. and J. Raboch, *Depression, antidepressants, and peripheral blood components.* Neuro Endocrinol Lett, 2008. **29**(1): p. 17-28.
39. Etkin, A., K.E. Prater, A.F. Schatzberg, V. Menon, and M.D. Greicius, *Disrupted amygdalar subregion functional connectivity and evidence of a compensatory network in generalized anxiety disorder.* Arch Gen Psychiatry, 2009. **66**(12): p. 1361-72.
40. Keedwell, P.A., D. Drapier, S. Surguladze, V. Giampietro, M. Brammer, and M. Phillips, *Subgenual cingulate and visual cortex responses to sad faces predict clinical outcome during antidepressant treatment for depression.* J Affect Disord, 2010. **120**(1-3): p. 120-5.
41. Pizzagalli, D.A., A.J. Holmes, D.G. Dillon, E.L. Goetz, J.L. Birk, R. Bogdan, D.D. Dougherty, D.V. Iosifescu, S.L. Rauch, and M. Fava, *Reduced caudate and nucleus accumbens response to rewards in unmedicated individuals with major depressive disorder.* Am J Psychiatry, 2009. **166**(6): p. 702-10.

42. Bruder, G.E., J.P. Sedoruk, J.W. Stewart, P.J. McGrath, F.M. Quitkin, and C.E. Tenke, *Electroencephalographic alpha measures predict therapeutic response to a selective serotonin reuptake inhibitor antidepressant: pre- and post-treatment findings.* Biol Psychiatry, 2008. **63**(12): p. 1171-7.
43. Iosifescu, D.V., S. Greenwald, P. Devlin, D. Mischoulon, J.W. Denninger, J.E. Alpert, and M. Fava, *Frontal EEG predictors of treatment outcome in major depressive disorder.* Eur Neuropsychopharmacol, 2009. **19**(11): p. 772-7.
44. Knott, V.J., J.I. Telner, Y.D. Lapierre, M. Browne, and E.R. Horn, *Quantitative EEG in the prediction of antidepressant response to imipramine.* J Affect Disord, 1996. **39**(3): p. 175-84.
45. Tenke, C.E., J. Kayser, N.A. Gates, D.M. Alschuler, C.J. Kroppmann, S. Fekri, C.G. Manna, J.W. Stewart, P.J. McGrath, and G.E. Bruder, *Auditory evoked potential (AEP) and EEG measures in depressed patients predict response to antidepressants.* Biol. Psychiatry, 2010. **67**(1s-271s): p. 98s.
46. Ulrich, G., E. Renfordt, and K. Frick, *The topographical distribution of alpha-activity in the resting eeg of endogenous-depressive inpatients with and without clinical-response to pharmacotherapy.* Pharmacopsychiatry, 1986. **19**: p. 272-273.
47. Korb, A.S., A.M. Hunter, I.A. Cook, and A.F. Leuchter, *Rostral anterior cingulate cortex theta current density and response to antidepressants and placebo in major depression.* Clin Neurophysiol, 2009. **120**(7): p. 1313-9.
48. Mulert, C., G. Juckel, M. Brunnmeier, S. Karch, G. Leicht, R. Mergl, H.J. Moller, U. Hegerl, and O. Pogarell, *Prediction of treatment response in major depression: integration of concepts.* J Affect Disord, 2007. **98**(3): p. 215-25.
49. Pizzagalli, D., R.D. Pascual-Marqui, J.B. Nitschke, T.R. Oakes, C.L. Larson, H.C. Abercrombie, S.M. Schaefer, J.V. Koger, R.M. Benca, and R.J. Davidson, *Anterior cingulate activity as a predictor of degree of treatment response in major depression: evidence from brain electrical tomography analysis.* Am J Psychiatry, 2001. **158**(3): p. 405-15.
50. Gallinat, J., R. Bottlender, G. Juckel, A. Munke-Puchner, G. Stotz, H.J. Kuss, P. Mavrogiorgou, and U. Hegerl, *The loudness dependency of the auditory evoked N1/P2-component as a predictor of the acute SSRI response in depression.* Psychopharmacology (Berl), 2000. **148**(4): p. 404-11.
51. Hegerl, U., G. Juckel, and H.J. Moller, *[Event related brain potentials as indicators of neurochemical dysfunctions in psychiatric patients].* Nervenarzt, 1996. **67**(5): p. 360-8.
52. Paige, S.R., D.F. Fitzpatrick, J.P. Kline, S.E. Balogh, and S.E. Hendricks, *Event-related potential amplitude/intensity slopes predict response to antidepressants.* Neuropsychobiology, 1994. **30**(4): p. 197-201.
53. Holsboer, F., *How can we realize the promise of personalized antidepressant medicines?* Nat Rev Neurosci, 2008. **9**(8): p. 638-46.
54. Frank, E., D.J. Kupfer, E.F. Wagner, A.B. McEachran, and C. Cornes, *Efficacy of interpersonal psychotherapy as a maintenance treatment of recurrent depression. Contributing factors.* Arch Gen Psychiatry, 1991. **48**(12): p. 1053-9.
55. Helfand, M., D.I. Buckley, M. Freeman, R. Fu, K. Rogers, C. Fleming, and L.L. Humphrey, *Emerging risk factors for coronary heart disease: a summary of systematic reviews conducted for the U.S. Preventive Services Task Force.* Ann Intern Med, 2009. **151**(7): p. 496-507.
56. Ustun, T.B., *The global burden of mental disorders.* Am J Public Health, 1999. **89**(9): p. 1315-8.
57. Murray, C.J. and A.D. Lopez, *Evidence-based health policy--lessons from the Global Burden of Disease Study.* Science, 1996. **274**(5288): p. 740-3.
58. Kessler, R.C., P. Berglund, O. Demler, R. Jin, D. Koretz, K.R. Merikangas, A.J. Rush, E.E. Walters, P.S. Wang, and R. National Comorbidity Survey, *The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R).[see comment].* Journal of the American Medical Association, 2003. **289**(23): p. 3095-105.
59. Rush, A.J., H.C. Kraemer, H.A. Sackeim, M. Fava, M.H. Trivedi, E. Frank, P.T. Ninan, M.E. Thase, A.J. Gelenberg, D.J. Kupfer, D.A. Regier, J.F. Rosenbaum, O. Ray, and A.F. Schatzberg, *Report by the ACNP Task Force on response and remission in major depressive disorder.* Neuropsychopharmacology, 2006. **31**(9): p. 1841-53.

60. Conn, P.J. and B.L. Roth, *Opportunities and challenges of psychiatric drug discovery: roles for scientists in academic, industry, and government settings.* Neuropsychopharmacology, 2008. **33**(9): p. 2048-60.
61. Hasler, G., W.C. Drevets, H.K. Manji, and D.S. Charney, *Discovering endophenotypes for major depression.* Neuropsychopharmacology, 2004. **29**(10): p. 1765-81.
62. Insel, T.R. and P.S. Wang, *The STAR*D trial: revealing the need for better treatments.* Psychiatr Serv, 2009. **60**(11): p. 1466-7.
63. Kornstein, S.G., A.F. Schatzberg, M.E. Thase, K.A. Yonkers, J.P. McCullough, G.I. Keitner, A.J. Gelenberg, S.M. Davis, W.M. Harrison, and M.B. Keller, *Gender differences in treatment response to sertraline versus imipramine in chronic depression.* Am J Psychiatry, 2000. **157**(9): p. 1445-52.
64. Papakostas, G.I., S.M. Stahl, A. Krishen, C.A. Seifert, V.L. Tucker, E.P. Goodale, and M. Fava, *Efficacy of bupropion and the selective serotonin reuptake inhibitors in the treatment of major depressive disorder with high levels of anxiety (anxious depression): a pooled analysis of 10 studies.* J Clin Psychiatry, 2008. **69**(8): p. 1287-92.
65. Coryell, W., *The treatment of psychotic depression.* J Clin Psychiatry, 1998. **59 Suppl 1**: p. 22-7; discussion 28-9.
66. McGrath, P.J., J.W. Stewart, E.V. Nunes, K. Ocepek-Welikson, J.G. Rabkin, F.M. Quitkin, and D.F. Klein, *A double-blind crossover trial of imipramine and phenelzine for outpatients with treatment-refractory depression.* Am J Psychiatry, 1993. **150**(1): p. 118-23.
67. Quitkin, F.M., W. Harrison, J.W. Stewart, P.J. McGrath, E. Tricamo, K. Ocepek-Welikson, J.G. Rabkin, S.G. Wager, E. Nunes, and D.F. Klein, *Response to phenelzine and imipramine in placebo nonresponders with atypical depression. A new application of the crossover design.* Arch Gen Psychiatry, 1991. **48**(4): p. 319-23.
68. Quitkin, F.M., P.J. McGrath, J.W. Stewart, W. Harrison, S.G. Wager, E. Nunes, J.G. Rabkin, E. Tricamo, J. Markowitz, and D.F. Klein, *Phenelzine and imipramine in mood reactive depressives. Further delineation of the syndrome of atypical depression.* Arch Gen Psychiatry, 1989. **46**(9): p. 787-93.
69. Fournier, J.C., R.J. DeRubeis, R.C. Shelton, S.D. Hollon, J.D. Amsterdam, and R. Gallop, *Prediction of response to medication and cognitive therapy in the treatment of moderate to severe depression.* J Consult Clin Psychol, 2009. **77**(4): p. 775-87.
70. Lafer, B., A.A. Nierenberg, J.F. Rosenbaum, and M. Fava, *Outpatients with DSM-III-R versus DSM-IV melancholic depression.* Compr Psychiatry, 1996. **37**(1): p. 37-9.
71. Fava, M., L.A. Uebelacker, J.E. Alpert, A.A. Nierenberg, J.A. Pava, and J.F. Rosenbaum, *Major depressive subtypes and treatment response.* Biol Psychiatry, 1997. **42**(7): p. 568-76.
72. Golden, C., *Stroop Color and Word Test: A Manual for Clinical and Experimental Uses*. 1978, Wood Dale, IL: Stoelting Co.
73. Tenke, C.E., J. Kayser, N.A. Gates, D.M. Alschuler, C.J. Kroppmann, S. Fekri, J.W. Stewart, P.J. McGrath, and G.E. Bruder, *Characterization of N1/P2 loudness dependency by temporal principal components analysis of current source density (CSD-PCA): prediction of treatment response in depressed patients.* Psychophysiology, 2009. **46**: p. S39-S40.
74. Fava, M., R.D. Vuolo, E.C. Wright, A.A. Nierenberg, J.E. Alpert, and J.F. Rosenbaum, *Fenfluramine challenge in unipolar depression with and without anger attacks.* Psychiatry Res, 2000. **94**(1): p. 9-18.
75. Redolat, R., M.C. Gomez, P. Vicens, and M.C. Carrasco, *Bupropion effects on aggressiveness and anxiety in OF1 male mice.* Psychopharmacology (Berl), 2005. **177**(4): p. 418-27.
76. Damluji, N.F. and J.M. Ferguson, *Paradoxical worsening of depressive symptomatology caused by antidepressants.* J Clin Psychopharmacol, 1988. **8**(5): p. 347-9.
77. Butler, A.C., J.E. Chapman, E.M. Forman, and A.T. Beck, *The empirical status of cognitive-behavioral therapy: a review of meta-analyses.* Clin Psychol Rev, 2006. **26**(1): p. 17-31.
78. Fava, M., *Daytime sleepiness and insomnia as correlates of depression.* J Clin Psychiatry, 2004. **65 Suppl 16**: p. 27-32.
79. Papakostas, G.I., D.J. Nutt, L.A. Hallett, V.L. Tucker, A. Krishen, and M. Fava, *Resolution of sleepiness and fatigue in major depressive disorder: A comparison of bupropion and the selective serotonin reuptake inhibitors.* Biol Psychiatry, 2006. **60**(12): p. 1350-5.

80. Bruder, G.E., J.W. Stewart, C.E. Tenke, P.J. McGrath, P. Leite, N. Bhattacharya, and F.M. Quitkin, *Electroencephalographic and perceptual asymmetry differences between responders and nonresponders to an SSRI antidepressant.* Biol Psychiatry, 2001. **49**(5): p. 416-25.
81. Kayser, J., *Polygraphic Recording Data Exchange - PolyRex (http://psychophysiology.cpmc.columbia.edu/PolyRex.htm)*. 2003: New York State Psychiatric Institute: Department of Biopsychology.
82. Kayser, J. and C.E. Tenke, *Optimizing PCA methodology for ERP component identification and measurement: theoretical rationale and empirical evaluation.* Clin Neurophysiol, 2003. **114**(12): p. 2307-25.
83. Kayser, J. and C.E. Tenke, *Principal components analysis of Laplacian waveforms as a generic method for identifying ERP generator patterns: II. Adequacy of low-density estimates.* Clin Neurophysiol, 2006. **117**(2): p. 369-80.
84. Kayser, J. and C.E. Tenke, *Trusting in or breaking with convention: towards a renaissance of principal components analysis in electrophysiology.* Clin Neurophysiol, 2005. **116**(8): p. 1747-53.
85. Taylor, B.P., G.E. Bruder, J.W. Stewart, P.J. McGrath, J. Halperin, H. Ehrlichman, and F.M. Quitkin, *Psychomotor slowing as a predictor of fluoxetine nonresponse in depressed outpatients.* Am J Psychiatry, 2006. **163**(1): p. 73-8.
86. Tombaugh, T., J. Kozak, and L. Rees, *Normative Data for the Controlled Oral Association Test*, in *A Compendium of Neuropsychological Tests. 2nd ed.*, O. Spreen and E. Strauss, Editors. 1998, Oxford University Press: New York.
87. Quitkin, F., A. Rifkin, and D.F. Klein, *Monoamine oxidase inhibitors. A review of antidepressant effectiveness.* Arch Gen Psychiatry, 1979. **36**(7): p. 749-60.
88. Quitkin, F.M., W. Harrison, M. Liebowitz, P. McGrath, J.G. Rabkin, J. Stewart, and J. Markowitz, *Defining the boundaries of atypical depression.* J Clin Psychiatry, 1984. **45**(7 Pt 2): p. 19-21.
89. Rinne, T., W. van den Brink, L. Wouters, and R. van Dyck, *SSRI treatment of borderline personality disorder: a randomized, placebo-controlled clinical trial for female patients with borderline personality disorder.* Am J Psychiatry, 2002. **159**(12): p. 2048-54.
90. Joffe, R.T., R.M. Bagby, and A. Levitt, *Anxious and nonanxious depression.* Am J Psychiatry, 1993. **150**(8): p. 1257-8.
91. Clayton, P.J., W.M. Grove, W. Coryell, M. Keller, R. Hirschfeld, and J. Fawcett, *Follow-up and family study of anxious depression.* Am J Psychiatry, 1991. **148**(11): p. 1512-7.
92. Insel, T.R. and B.N. Cuthbert, *Endophenotypes: bridging genomic complexity and disorder heterogeneity.* Biol. Psychiatry, 2009. **66**(11): p. 988-9.
93. Bertolino, A., G. Arciero, V. Rubino, V. Latorre, M. De Candia, V. Mazzola, G. Blasi, G. Caforio, A. Hariri, B. Kolachana, M. Nardini, D.R. Weinberger, and T. Scarabino, *Variation of human amygdala response during threatening stimuli as a function of 5'HTTLPR genotype and personality style.* Biol Psychiatry, 2005. **57**(12): p. 1517-25.
94. Furmark, T., M. Tillfors, H. Garpenstrand, I. Marteinsdottir, B. Langstrom, L. Oreland, and M. Fredrikson, *Serotonin transporter polymorphism related to amygdala excitability and symptom severity in patients with social phobia.* Neurosci Lett, 2004. **362**(3): p. 189-92.
95. Hariri, A.R., E.M. Drabant, K.E. Munoz, B.S. Kolachana, V.S. Mattay, M.F. Egan, and D.R. Weinberger, *A susceptibility gene for affective disorders and the response of the human amygdala.* Arch Gen Psychiatry, 2005. **62**(2): p. 146-52.
96. Hariri, A.R., V.S. Mattay, A. Tessitore, B. Kolachana, F. Fera, D. Goldman, M.F. Egan, and D.R. Weinberger, *Serotonin transporter genetic variation and the response of the human amygdala.* Science, 2002. **297**(5580): p. 400-3.
97. Heinz, A., D.F. Braus, M.N. Smolka, J. Wrase, I. Puls, D. Hermann, S. Klein, S.M. Grusser, H. Flor, G. Schumann, K. Mann, and C. Buchel, *Amygdala-prefrontal coupling depends on a genetic variation of the serotonin transporter.* Nat Neurosci, 2005. **8**(1): p. 20-1.
98. Heinz, A., M.N. Smolka, D.F. Braus, J. Wrase, A. Beck, H. Flor, K. Mann, G. Schumann, C. Buchel, A.R. Hariri, and D.R. Weinberger, *Serotonin transporter genotype (5-HTTLPR): effects of neutral and undefined conditions on amygdala activation.* Biol Psychiatry, 2007. **61**(8): p. 1011-4.

99. Chen, C.H., J. Suckling, C. Ooi, C.H. Fu, S.C. Williams, N.D. Walsh, M.T. Mitterschiffthaler, E.M. Pich, and E. Bullmore, *Functional coupling of the amygdala in depressed patients treated with antidepressant medication.* Neuropsychopharmacology, 2008. **33**(8): p. 1909-18.
100. Fu, C.H., J. Mourao-Miranda, S.G. Costafreda, A. Khanna, A.F. Marquand, S.C. Williams, and M.J. Brammer, *Pattern classification of sad facial processing: toward the development of neurobiological markers in depression.* Biol Psychiatry, 2008. **63**(7): p. 656-62.
101. Mayberg, H.S., S.K. Brannan, R.K. Mahurin, P.A. Jerabek, J.S. Brickman, J.L. Tekell, J.A. Silva, S. McGinnis, T.G. Glass, C.C. Martin, and P.T. Fox, *Cingulate function in depression: a potential predictor of treatment response.* Neuroreport, 1997. **8**(4): p. 1057-61.
102. Siegle, G.J., C.S. Carter, and M.E. Thase, *Use of FMRI to predict recovery from unipolar depression with cognitive behavior therapy.* Am J Psychiatry, 2006. **163**(4): p. 735-8.
103. Chen, C.H., K. Ridler, J. Suckling, S. Williams, C.H. Fu, E. Merlo-Pich, and E. Bullmore, *Brain imaging correlates of depressive symptom severity and predictors of symptom improvement after antidepressant treatment.* Biol Psychiatry, 2007. **62**(5): p. 407-14.
104. Greicius, M.D., B.H. Flores, V. Menon, G.H. Glover, H.B. Solvason, H. Kenna, A.L. Reiss, and A.F. Schatzberg, *Resting-state functional connectivity in major depression: abnormally increased contributions from subgenual cingulate cortex and thalamus.* Biol Psychiatry, 2007. **62**(5): p. 429-37.
105. Anand, A., Y. Li, Y. Wang, K. Gardner, and M.J. Lowe, *Reciprocal effects of antidepressant treatment on activity and connectivity of the mood regulating circuit: an FMRI study.* J Neuropsychiatry Clin Neurosci, 2007. **19**(3): p. 274-82.
106. Anand, A., Y. Li, Y. Wang, J. Wu, S. Gao, L. Bukhari, V.P. Mathews, A. Kalnin, and M.J. Lowe, *Activity and connectivity of brain mood regulating circuit in depression: a functional magnetic resonance study.* Biol Psychiatry, 2005. **57**(10): p. 1079-88.
107. Sheline, Y.I., D.M. Barch, J.L. Price, M.M. Rundle, S.N. Vaishnavi, A.Z. Snyder, M.A. Mintun, S. Wang, R.S. Coalson, and M.E. Raichle, *The default mode network and self-referential processes in depression.* Proc Natl Acad Sci U S A, 2009. **106**(6): p. 1942-7.
108. Cardinal, R.N., C.A. Winstanley, T.W. Robbins, and B.J. Everitt, *Limbic corticostriatal systems and delayed reinforcement.* Ann N Y Acad Sci, 2004. **1021**: p. 33-50.
109. Parkinson, J.A., J.W. Dalley, R.N. Cardinal, A. Bamford, B. Fehnert, G. Lachenal, N. Rudarakanchana, K.M. Halkerston, T.W. Robbins, and B.J. Everitt, *Nucleus accumbens dopamine depletion impairs both acquisition and performance of appetitive Pavlovian approach behaviour: implications for mesoaccumbens dopamine function.* Behav Brain Res, 2002. **137**(1-2): p. 149-63.
110. Schultz, W., *Dopamine neurons and their role in reward mechanisms.* Curr Opin Neurobiol, 1997. **7**(2): p. 191-7.
111. Schultz, W., *Behavioral dopamine signals.* Trends Neurosci, 2007. **30**(5): p. 203-10.
112. Knutson, B. and J.C. Cooper, *Functional magnetic resonance imaging of reward prediction.* Curr Opin Neurol, 2005. **18**(4): p. 411-7.
113. Epstein, J., H. Pan, J.H. Kocsis, Y. Yang, T. Butler, J. Chusid, H. Hochberg, J. Murrough, E. Strohmayer, E. Stern, and D.A. Silbersweig, *Lack of ventral striatal response to positive stimuli in depressed versus normal subjects.* Am J Psychiatry, 2006. **163**(10): p. 1784-90.
114. Mori, S. and J. Zhang, *Principles of diffusion tensor imaging and its applications to basic neuroscience research.* Neuron, 2006. **51**(5): p. 527-39.
115. Sexton, C.E., C.E. Mackay, and K.P. Ebmeier, *A systematic review of diffusion tensor imaging studies in affective disorders.* Biol Psychiatry, 2009. **66**(9): p. 814-23.
116. Taylor, W.D., M. Kuchibhatla, M.E. Payne, J.R. Macfall, Y.I. Sheline, K.R. Krishnan, and P.M. Doraiswamy, *Frontal white matter anisotropy and antidepressant remission in late-life depression.* PLoS One, 2008. **3**(9): p. e3267.
117. Fischl, B. and A.M. Dale, *Measuring the thickness of the human cerebral cortex from magnetic resonance images.* Proc Natl Acad Sci U S A, 2000. **97**(20): p. 11050-5.
118. Peterson, B.S., V. Warner, R. Bansal, H. Zhu, X. Hao, J. Liu, K. Durkin, P.B. Adams, P. Wickramaratne, and M.M. Weissman, *Cortical thinning in persons at increased familial risk for major depression.* Proc Natl Acad Sci U S A, 2009. **106**(15): p. 6273-8.

119. Arango, V., M.D. Underwood, A.V. Gubbi, and J.J. Mann, *Localized alterations in pre- and postsynaptic serotonin binding sites in the ventrolateral prefrontal cortex of suicide victims.* Brain Res, 1995. **688**(1-2): p. 121-33.
120. Leuchter, A.F., I.A. Cook, W.S. Gilmer, L.B. Marangell, K.S. Burgoyne, R.H. Howland, M.H. Trivedi, S. Zisook, R. Jain, M. Fava, D. Iosifescu, and S. Greenwald, *Effectiveness of a quantitative electroencephalographic biomarker for predicting differential response or remission with escitalopram and bupropion in major depressive disorder.* Psychiatry Res, 2009. **169**(2): p. 132-8.
121. Ulrich, G., Renfordt, E., & Frick, K., *The topographical distribution of alpha-activity in the resting eeg of endogenous-depressive inpatients with and without clinical-response to pharmacotherapy.* Pharmacopsychiatry, 1986(19): p. 272-273.
122. Little, J.T., T.A. Ketter, T.A. Kimbrell, R.T. Dunn, B.E. Benson, M.W. Willis, D.A. Luckenbaugh, and R.M. Post, *Bupropion and venlafaxine responders differ in pretreatment regional cerebral metabolism in unipolar depression.* Biol Psychiatry, 2005. **57**(3): p. 220-8.
123. Hegerl, U. and G. Juckel, *Intensity dependence of auditory evoked potentials as an indicator of central serotonergic neurotransmission: a new hypothesis.* Biol Psychiatry, 1993. **33**(3): p. 173-87.
124. Lewis, D.A., M.J. Campbell, R.D. Terry, and J.H. Morrison, *Laminar and regional distributions of neurofibrillary tangles and neuritic plaques in Alzheimer's disease: a quantitative study of visual and auditory cortices.* J Neurosci, 1987. **7**(6): p. 1799-808.
125. Campbell, M.J., D.A. Lewis, S.L. Foote, and J.H. Morrison, *Distribution of choline acetyltransferase-, serotonin-, dopamine-beta-hydroxylase-, tyrosine hydroxylase-immunoreactive fibers in monkey primary auditory cortex.* J Comp Neurol, 1987. **261**(2): p. 209-20.
126. Naatanen, R. and T. Picton, *The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure.* Psychophysiology, 1987. **24**(4): p. 375-425.
127. Juckel, G., U. Hegerl, M. Molnar, V. Csepe, and G. Karmos, *Auditory evoked potentials reflect serotonergic neuronal activity--a study in behaving cats administered drugs acting on 5-HT1A autoreceptors in the dorsal raphe nucleus.* Neuropsychopharmacology, 1999. **21**(6): p. 710-6.
128. Linka, T., B.W. Muller, S. Bender, G. Sartory, and M. Gastpar, *The intensity dependence of auditory evoked ERP components predicts responsiveness to reboxetine treatment in major depression.* Pharmacopsychiatry, 2005. **38**(3): p. 139-43.
129. Flament, M.F., R.M. Lane, R. Zhu, and Z. Ying, *Predictors of an acute antidepressant response to fluoxetine and sertraline.* Int Clin Psychopharmacol, 1999. **14**(5): p. 259-75.
130. Kalayam, B. and G.S. Alexopoulos, *Prefrontal dysfunction and treatment response in geriatric depression.* Arch Gen Psychiatry, 1999. **56**(8): p. 713-8.
131. Caligiuri, M.P., V. Gentili, S. Eberson, J. Kelsoe, M. Rapaport, and J.C. Gillin, *A quantitative neuromotor predictor of antidepressant non-response in patients with major depression.* J Affect Disord, 2003. **77**(2): p. 135-41.
132. Rampello, L., G. Nicoletti, and R. Raffaele, *Dopaminergic hypothesis for retarded depression: a symptom profile for predicting therapeutical responses.* Acta Psychiatr Scand, 1991. **84**(6): p. 552-4.
133. Herrera-Guzman, I., E. Gudayol-Ferre, J. Lira-Mandujano, J. Herrera-Abarca, D. Herrera-Guzman, K. Montoya-Perez, and J. Guardia-Olmos, *Cognitive predictors of treatment response to bupropion and cognitive effects of bupropion in patients with major depressive disorder.* Psychiatry Res, 2008. **160**(1): p. 72-82.
134. Kennedy, S.H., K.R. Evans, S. Kruger, H.S. Mayberg, J.H. Meyer, S. McCann, A.I. Arifuzzman, S. Houle, and F.J. Vaccarino, *Changes in regional brain glucose metabolism measured with positron emission tomography after paroxetine treatment of major depression.* Am J Psychiatry, 2001. **158**(6): p. 899-905.
135. Saxena, S., A.L. Brody, M.L. Ho, N. Zohrabi, K.M. Maidment, and L.R. Baxter, Jr., *Differential brain metabolic predictors of response to paroxetine in obsessive-compulsive disorder versus major depression.* Am J Psychiatry, 2003. **160**(3): p. 522-32.
136. Langenecker, S.A., S.E. Kennedy, L.M. Guidotti, E.M. Briceno, L.S. Own, T. Hooven, E.A. Young, H. Akil, D.C. Noll, and J.K. Zubieta, *Frontal and limbic activation during inhibitory control predicts treatment response in major depressive disorder.* Biol Psychiatry, 2007. **62**(11): p. 1272-80.

137. Pizzagalli, D.A., L.A. Peccoralo, R.J. Davidson, and J.D. Cohen, *Resting anterior cingulate activity and abnormal responses to errors in subjects with elevated depressive symptoms: a 128-channel EEG study.* Hum Brain Mapp, 2006. **27**(3): p. 185-201.
138. Holmes, A.J. and D.A. Pizzagalli, *Spatiotemporal dynamics of error processing dysfunctions in major depressive disorder.* Arch Gen Psychiatry, 2008. **65**(2): p. 179-88.
139. Holmes, A.J., R. Bogdan, and D.A. Pizzagalli, *Serotonin transporter genotype and action monitoring dysfunction: a possible substrate underlying increased vulnerability to depression.* Neuropsychopharmacology, 2010. **35**(5): p. 1186-97.
140. Julian, L.J. and D.C. Mohr, *Cognitive predictors of response to treatment for depression in multiple sclerosis.* J Neuropsychiatry Clin Neurosci, 2006. **18**(3): p. 356-63.
141. Gorlyn, M., J.G. Keilp, M.F. Grunebaum, B.P. Taylor, M.A. Oquendo, G.E. Bruder, J.W. Stewart, G. Zalsman, and J.J. Mann, *Neuropsychological characteristics as predictors of SSRI treatment response in depressed subjects.* J Neural Transm, 2008. **115**(8): p. 1213-9.
142. Herrera-Guzman, I., E. Gudayol-Ferre, D. Herrera-Guzman, J. Guardia-Olmos, E. Hinojosa-Calvo, and J.E. Herrera-Abarca, *Effects of selective serotonin reuptake and dual serotonergic-noradrenergic reuptake treatments on memory and mental processing speed in patients with major depressive disorder.* J Psychiatr Res, 2009. **43**(9): p. 855-63.
143. Hundt, N.E., R.O. Nelson-Gray, N.A. Kimbrel, J.T. Mitchell, and T.R. Kwapil, *The interaction of reinforcement sensitivity and life events in the prediction of anhedonic depression and mixed anxiety-depression symptoms.* Personality and Individual Differences, 2007. **43**: p. 1001-1012.
144. McFarland, B.R., S.A. Shankman, C.E. Tenke, G.E. Bruder, and D.N. Klein, *Behavioral activation system deficits predict the six-month course of depression.* J Affect Disord, 2006. **91**(2-3): p. 229-34.
145. Kasch, K.L., J. Rottenberg, B.A. Arnow, and I.H. Gotlib, *Behavioral activation and inhibition systems and the severity and course of depression.* J Abnorm Psychol, 2002. **111**(4): p. 589-97.
146. Spijker, J., R.V. Bijl, R. de Graaf, and W.A. Nolen, *Determinants of poor 1-year outcome of DSM-III-R major depression in the general population: results of the Netherlands Mental Health Survey and Incidence Study (NEMESIS).* Acta Psychiatr Scand, 2001. **103**(2): p. 122-30.
147. Moos, R.H. and R.C. Cronkite, *Symptom-based predictors of a 10-year chronic course of treated depression.* J Nerv Ment Dis, 1999. **187**(6): p. 360-8.
148. Pizzagalli, D.A., D. Iosifescu, L.A. Hallett, K.G. Ratner, and M. Fava, *Reduced hedonic capacity in major depressive disorder: evidence from a probabilistic reward task.* J Psychiatr Res, 2008. **43**(1): p. 76-87.
149. Pizzagalli, D.A., A.L. Jahn, and J.P. O'Shea, *Toward an objective characterization of an anhedonic phenotype: a signal-detection approach.* Biol Psychiatry, 2005. **57**(4): p. 319-27.
150. Pizzagalli, D.A., A.E. Evins, E.C. Schetter, M.J. Frank, P.E. Pajtas, D.L. Santesso, and M. Culhane, *Single dose of a dopamine agonist impairs reinforcement learning in humans: behavioral evidence from a laboratory-based measure of reward responsiveness.* Psychopharmacology (Berl), 2008. **196**(2): p. 221-32.
151. Santesso, D.L., D.G. Dillon, J.L. Birk, A.J. Holmes, E. Goetz, R. Bogdan, and D.A. Pizzagalli, *Individual differences in reinforcement learning: behavioral, electrophysiological, and neuroimaging correlates.* Neuroimage, 2008. **42**(2): p. 807-16.
152. Bruheim, S., Y. Xi, J. Ju, and O. Fodstad, *Gene expression profiles classify human osteosarcoma xenografts according to sensitivity to doxorubicin, cisplatin, and ifosfamide.* Clin Cancer Res, 2009. **15**(23): p. 7161-9.
153. Hennings, J.M., T. Owashi, E.B. Binder, S. Horstmann, A. Menke, S. Kloiber, T. Dose, B. Wollweber, D. Spieler, T. Messer, R. Lutz, H. Kunzel, T. Bierner, T. Pollmacher, H. Pfister, T. Nickel, A. Sonntag, M. Uhr, M. Ising, F. Holsboer, and S. Lucae, *Clinical characteristics and treatment outcome in a representative sample of depressed inpatients - findings from the Munich Antidepressant Response Signature (MARS) project.* J Psychiatr Res, 2009. **43**(3): p. 215-29.
154. Uher, R., P. Huezo-Diaz, N. Perroud, R. Smith, M. Rietschel, O. Mors, J. Hauser, W. Maier, D. Kozel, N. Henigsberg, M. Barreto, A. Placentino, M.Z. Dernovsek, T.G. Schulze, P. Kalember, A. Zobel, P.M. Czerski, E.R. Larsen, D. Souery, C. Giovannini, J.M. Gray, C.M. Lewis, A. Farmer, K.J. Aitchison, P. McGuffin, and I. Craig, *Genetic predictors of response to antidepressants in the GENDEP project.* Pharmacogenomics J, 2009.

155. Montgomery, S.A. and M. Asberg, *A new depression scale designed to be sensitive to change.* Br J Psychiatry, 1979. **134**: p. 382-9.
156. Hamilton, M., *A rating scale for depression.* Journal of Neurology, Neurosurgery, & Psychiatry, 1960. **23**: p. 56-61.
157. Hamilton, M., *Development of a rating scale for primary depressive illness.* Br J Soc Clin Psychol, 1967. **6**(4): p. 278-96.
158. American Psychiatric Association, *Diagnostic and Statistical Manual of Mental Disorders - Text Revision.* 2000, Washington, DC: American Psychiatric Press.
159. Stewart, J.W., P.J. McGrath, F.M. Quitkin, W. Harrison, J. Markowitz, S. Wager, and M.R. Leibowitz, *Relevance of DMS-III depressive subtype and chronicity of antidepressant efficacy in atypical depression. Differential response to phenelzine, imipramine, and placebo.* Arch Gen Psychiatry, 1989. **46**(12): p. 1080-7.
160. Rush, A.J., M.H. Trivedi, H.M. Ibrahim, T.J. Carmody, B. Arnow, D.N. Klein, J.C. Markowitz, P.T. Ninan, S. Kornstein, R. Manber, M.E. Thase, J.H. Kocsis, and M.B. Keller, *The 16-item Quick Inventory of Depressive Symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): A psychometric evaluation in patients with chronic major depression.* Biological Psychiatry, 2003. **54**(5): p. 573-583.
161. Trivedi, M.H., A.J. Rush, H.M. Ibrahim, T.J. Carmody, M.M. Biggs, T. Suppes, M.L. Crismon, K. Shores-Wilson, M.G. Toprac, E.B. Dennehy, B. Witte, and T.M. Kashner, *The Inventory of Depressive Symptomatology, Clinician Rating (IDS-C) and Self-Report (IDS-SR), and the Quick Inventory of Depressive Symptomatology, Clinician Rating (QIDS-C) and Self-Report (QIDS-SR) in public sector patients with mood disorders: a psychometric evaluation.* Psychological Medicine, 2004. **34**(1): p. 73-82.
162. Wisniewski, S.R., A.J. Rush, G.K. Balasubramani, M.H. Trivedi, A. Nierenberg, and for the STAR*D Investigators, *Self-Rated Global Measure of the Frequency, Intensity and Burden of Medication Side Effects.* Journal of Psychiatric Practice, 2006. **12**(2): p. 71-79.
163. Fava, M., *Diagnosis and definition of treatment-resistant depression.* Biol Psychiatry, 2003. **53**(8): p. 649-59.
164. Moran, P., M. Leese, T. Lee, P. Walters, G. Thornicroft, and A. Mann, *Standardised Assessment of Personality - Abbreviated Scale (SAPAS): preliminary validation of a brief screen for personality disorder.* Br J Psychiatry, 2003. **183**: p. 228-32.
165. Costa, P.T. and R.R. McCrae, *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual*. 1992, Odessa, FL: Psychological Assessment Resources.
166. Williams, J.B., *A structured interview guide for the Hamilton Depression Rating Scale.* Arch Gen Psychiatry, 1988. **45**(8): p. 742-7.
167. Posner, K., M.A. Oquendo, M. Gould, B. Stanley, and M. Davies, *Columbia Classification Algorithm of Suicide Assessment (C-CASA): classification of suicidal events in the FDA's pediatric suicidal risk analysis of antidepressants.* Am J Psychiatry, 2007. **164**(7): p. 1035-43.
168. Bernstein, D.P., J.A. Stein, M.D. Newcomb, E. Walker, D. Pogge, T. Ahluvalia, J. Stokes, L. Handelsman, M. Medrano, D. Desmond, and W. Zule, *Development and validation of a brief screening version of the Childhood Trauma Questionnaire.* Child Abuse Negl, 2003. **27**(2): p. 169-90.
169. Sangha, O., G. Stucki, M.H. Liang, A.H. Fossel, and J.N. Katz, *The Self-Administered Comorbidity Questionnaire: a new method to assess comorbidity for clinical and health services research.* Arthritis Rheum, 2003. **49**(2): p. 156-63.
170. Weissman, M.M., P. Wickramaratne, P. Adams, S. Wolk, H. Verdeli, and M. Olfson, *Brief screening for family psychiatric history: the family history screen.* Arch Gen Psychiatry, 2000. **57**(7): p. 675-82.
171. Rahe, R.H., J.D. McKean, Jr., and R.J. Arthur, *A longitudinal study of life-change and illness patterns.* J Psychosom Res, 1967. **10**(4): p. 355-66.
172. Altman, E.G., D. Hedeker, J.L. Peterson, and J.M. Davis, *The Altman Self-Rating Mania Scale.* Biol. Psychiatry, 1997. **42**(10): p. 948-55.
173. Miller, M.A. and R.H. Rahe, *Life changes scaling for the 1990s.* J Psychosom Res, 1997. **43**(3): p. 279-92.
174. Levine, J. and N.R. Schooler, *SAFTEE: a technique for the systematic assessment of side effects in clinical trials.* Psychopharmacol Bull, 1986. **22**(2): p. 343-81.

175. Labbate, L.A. and S.B. Lare, *Sexual dysfunction in male psychiatric outpatients: validity of the Massachusetts General Hospital Sexual Functioning Questionnaire.* Psychother Psychosom, 2001. **70**(4): p. 221-5.
176. Fava, M., J.F. Rosenbaum, M. McCarthy, J. Pava, R. Steingard, and E. Bless, *Anger attacks in depressed outpatients and their response to fluoxetine.* Psychopharmacol Bull, 1991. **27**(3): p. 275-9.
177. Winkler, D., E. Pjrek, J. Kindler, A. Heiden, and S. Kasper, *Validation of a simplified definition of anger attacks.* Psychother Psychosom, 2006. **75**(2): p. 103-6.
178. Stewart, J.W., P.J. McGrath, F.M. Quitkin, J.G. Rabkin, W. Harrison, S. Wager, E. Nunes, K. Ocepek-Welikson, and E. Tricamo, *Chronic depression: response to placebo, imipramine, and phenelzine.* J Clin Psychopharmacol, 1993. **13**(6): p. 391-6.
179. Hirschfeld, R.M., J.B. Williams, R.L. Spitzer, J.R. Calabrese, L. Flynn, P.E. Keck, Jr., L. Lewis, S.L. McElroy, R.M. Post, D.J. Rapport, J.M. Russell, G.S. Sachs, and J. Zajecka, *Development and validation of a screening instrument for bipolar spectrum disorder: the Mood Disorder Questionnaire.* Am J Psychiatry, 2000. **157**(11): p. 1873-5.
180. Weissman, M.M. and S. Bothwell, *Assessment of social adjustment by patient self-report.* Arch Gen Psychiatry, 1976. **33**(9): p. 1111-5.
181. Egner, T., A. Etkin, S. Gale, and J. Hirsch, *Dissociable neural systems resolve conflict from emotional versus nonemotional distracters.* Cereb Cortex, 2008. **18**(6): p. 1475-84.
182. Etkin, A., T. Egner, D.M. Peraza, E.R. Kandel, and J. Hirsch, *Resolving emotional conflict: a role for the rostral anterior cingulate cortex in modulating activity in the amygdala.* Neuron, 2006. **51**(6): p. 871-82.
183. Etkin, A., K.E. Prater, F. Hoeft, V. Menon, and A.F. Schatzberg, *Failure of anterior cingulate activation and connectivity with the amygdala during implicit regulation of emotional processing in generalized anxiety disorder.* American Journal of Psychiatry, 2010. **In press**.
184. Forbes, E.E., *Where's the fun in that? Broadening the focus on reward function in depression.* Biol Psychiatry, 2009. **66**(3): p. 199-200.
185. Van Dijk, K.R., T. Hedden, A. Venkataraman, K.C. Evans, S.W. Lazar, and R.L. Buckner, *Intrinsic functional connectivity as a tool for human connectomics: theory, properties, and optimization.* J Neurophysiol, 2009. **Epub ahead of print 11/06/09 DOI 10.1152/jn.00783.2009**.
186. Liu, H., S.M. Stufflebeam, J. Sepulcre, T. Hedden, and R.L. Buckner, *Evidence from intrinsic activity that asymmetry of the human brain is controlled by multiple factors.* Proc Natl Acad Sci U S A, 2009. **106**(48): p. 20499-503.
187. Gholipour, A., N. Kehtarnavaz, K. Gopinath, R. Briggs, and I. Panahi, *Average field map image template for Echo-Planar image analysis.* Conf Proc IEEE Eng Med Biol Soc, 2008. **2008**: p. 94-7.
188. Yushkevich, P.A., B.B. Avants, J. Pluta, S. Das, D. Minkoff, D. Mechanic-Hamilton, S. Glynn, S. Pickup, W. Liu, J.C. Gee, M. Grossman, and J.A. Detre, *A high-resolution computational atlas of the human hippocampus from postmortem magnetic resonance imaging at 9.4 T.* Neuroimage, 2009. **44**(2): p. 385-98.
189. Zhang, H., P.A. Yushkevich, D. Rueckert, and J.C. Gee, *Unbiased white matter atlas construction using diffusion tensor images.* Med Image Comput Comput Assist Interv, 2007. **10**(Pt 2): p. 211-8.
190. Segonne, F., A.M. Dale, E. Busa, M. Glessner, D. Salat, H.K. Hahn, and B. Fischl, *A hybrid approach to the skull stripping problem in MRI.* Neuroimage, 2004. **22**(3): p. 1060-75.
191. Marcus, D.S., T.R. Olsen, M. Ramaratnam, and R.L. Buckner, *The Extensible Neuroimaging Archive Toolkit: an informatics platform for managing, exploring, and sharing neuroimaging data.* Neuroinformatics, 2007. **5**(1): p. 11-34.
192. Marcus, D.S., T.H. Wang, J. Parker, J.G. Csernansky, J.C. Morris, and R.L. Buckner, *Open Access Series of Imaging Studies (OASIS): cross-sectional MRI data in young, middle aged, nondemented, and demented older adults.* J Cogn Neurosci, 2007. **19**(9): p. 1498-507.
193. Friedman, L. and G.H. Glover, *Reducing interscanner variability of activation in a multicenter fMRI study: controlling for signal-to-fluctuation-noise-ratio (SFNR) differences.* Neuroimage, 2006. **33**(2): p. 471-81.
194. Friedman, L., G.H. Glover, D. Krenz, and V. Magnotta, *Reducing inter-scanner variability of activation in a multicenter fMRI study: role of smoothness equalization.* Neuroimage, 2006. **32**(4): p. 1656-68.

195. Kayser, J. and C.E. Tenke, *Principal components analysis of Laplacian waveforms as a generic method for identifying ERP generator patterns: I. Evaluation with auditory oddball tasks.* Clin Neurophysiol, 2006. **117**(2): p. 348-68.
196. Tenke, C.E. and J. Kayser, *Reference-free quantification of EEG spectra: combining current source density (CSD) and frequency principal components analysis (fPCA).* Clin Neurophysiol, 2005. **116**(12): p. 2826-46.
197. Cook, I.A., A.F. Leuchter, E. Witte, M. Abrams, S.H. Uijtdehaage, W. Stubbeman, S. Rosenberg-Thompson, C. Anderson-Hanley, and J.J. Dunkin, *Neurophysiologic predictors of treatment response to fluoxetine in major depression.* Psychiatry Res, 1999. **85**(3): p. 263-73.
198. Pascual-Marqui, R.D., D. Lehmann, T. Koenig, K. Kochi, M.C. Merlo, D. Hell, and M. Koukkou, *Low resolution brain electromagnetic tomography (LORETA) functional imaging in acute, neuroleptic-naive, first-episode, productive schizophrenia.* Psychiatry Res, 1999. **90**(3): p. 169-79.
199. Talairach, J. and P. Tournoux, *Co-planar stereotaxic atlas of the human brain: 3-dimensional proportional system - an approach to cerebral imaging*. 1988, New York: Thieme Medical Publishers.
200. Towle, V.L., J. Bolanos, D. Suarez, K. Tan, R. Grzeszczuk, D.N. Levin, R. Cakmur, S.A. Frank, and J.P. Spire, *The spatial location of EEG electrodes: locating the best-fitting sphere relative to cortical anatomy.* Electroencephalogr Clin Neurophysiol, 1993. **86**(1): p. 1-6.
201. Thorne, D.R., S.G. Genser, H.C. Sing, and F.W. Hegge, *The Walter Reed performance assessment battery.* Neurobehav Toxicol Teratol, 1985. **7**(4): p. 415-8.
202. Benton, A.L., K. Hamsher, and A.B. Sivan, *Mulitlingual Aphasia Examination (3rd ed.)*. 1983, Iowa City: AJA Associates.
203. Baddeley, A.D., *A three minute reasoning test based on grammatical transformations.* Psychonomic Science, 1968. **10**: p. 1019-1025.
204. Laming, D., *Autocorrelation of choice-reaction times.* Acta Psychol (Amst), 1979. **43**(5): p. 381-412.
205. Rabbitt, P.M., *Errors and error correction in choice-response tasks.* J Exp Psychol, 1966. **71**(2): p. 264-72.
206. Warden, D., A.J. Rush, M. Trivedi, L. Ritz, D. Stegman, and S.R. Wisniewski, *Quality improvement methods as applied to a multicenter effectiveness trial--STAR*D.* Contemp.Clin.Trials, 2005. **26**(1): p. 95-112.
207. Petkova, E., T. Tarpey, and U. Govindarajulu, *Predicting potential placebo effect in drug treated subjects.* International Journal of Biostatistics, 2009. **5**(1): p. 5 DOI: 10.2202/1557-4679.1152.
208. Wold, H., *Soft modeling by latent variables: The nonlinear iterative partial least squares approach. In Perspectives in Probability and Statistics, Papers in Honour of M.S.* , ed. J.G. Bartlett. 1975, London: Academic Press.
209. McIntosh, A.R., F.L. Bookstein, J.V. Haxby, and C.L. Grady, *Spatial pattern analysis of functional brain images using partial least squares.* Neuroimage, 1996. **3**(3 Pt 1): p. 143-57.
210. McIntosh, A.R. and N.J. Lobaugh, *Partial least squares analysis of neuroimaging data: applications and advances.* Neuroimage, 2004. **23 Suppl 1**: p. S250-63.
211. Li, K.C., *Sliced inverse regression for dimension reduction. .* Journal of the American Statistical Association, 1991. **86**: p. 316-342.
212. Cook, R. and S. Weisburg, *Comment on 'Sliced inverse regression for dimension reduction' by K.C. Li.* Journal of the American Statistical Association, 1991. **86**: p. 328-332.
213. Xia, Y., H. Tong, W. Li, and L. Zhu, *An adaptive estimation of dimension reduction space. .* Journal of the Royal Statistical Society, Series B, 2002. **64**: p. 363-410.
214. Hyvarinen, J., E. Karhunen, and Oja, *Independent Component Anaysis*. 2001, New York: Wiley.
215. Duda, R., Hart, P., and Stork, D., *Pattern Classification, Second Edition*. 2000, New York: Wiley.
216. Ripley, B.D., *Pattern Recognition and Neural Networks*. 1996, Cambridge: Cambridge University Press.
217. Vapnik, V.N., *The Nature of Statistical Learning Theory.* 1995, Berlin. : Springer-Verlag.
218. Hastie, T., R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* Second Edition ed. 2009, New York.: Springer.
219. Golland, P., B. Fischl, M. Spiridon, N. Kanwisher, R.L. Buckner, M.E. Shenton, R. Kikinis, A. Dale, and W.E.L. Grim- son, *Discriminative analysis for image-based studies. .* Medical Image Computing and Computer- Assisted Intervention - MICCAI Proceedings, Part I, 2002. **2488**: p. 508-515.

220. Ford, J., H. Farid, F. Makedon, L.A. Flashman, T.W. McAllister, V. Megalooikonomou, and A.J. Saykin, *Patient classification of fMRI activation maps. .* Medical Image Computing and Computer-Assisted Intervention - MICCAI Proceedings, Part II 2003. **2879**: p. 58-65.
221. Mourao-Miranda, J., A.L. Bokde, C. Born, H. Hampel, and M. Stetter, *Classifying brain states and determining the discriminating activation patterns: Support Vector Machine on functional MRI data.* Neuroimage, 2005. **28**(4): p. 980-95.
222. Pessoa, L. and S. Padmala, *Decoding near-threshold perception of fear from distributed single-trial brain activation.* Cereb Cortex, 2007. **17**(3): p. 691-701.
223. Sato, J.R., A. Fujita, C.E. Thomaz, G. Martin Mda, J. Mourao-Miranda, M.J. Brammer, and E. Amaro Junior, *Evaluating SVM and MLDA in the extraction of discriminant regions for mental state prediction.* Neuroimage, 2009. **46**(1): p. 105-14.
224. Marquand, A.F., J. Mourao-Miranda, M.J. Brammer, A.J. Cleare, and C.H.Y. Fu, *Neuroanatomy of verbal working memory as a diagnostic biomarker for depression.* Neuroreport, 2008. **19**(15): p. 1507-11.
225. Breiman, L., J. Friedman, R. Olshen, and C. Stone, *Classification and Regression Trees*. 1984: Wadsworth.
226. Friedman, J., *Multivariate adaptive regression splines (with discussion).* Annals of Statistics, 1991. **19**(1): p. 1-141.
227. Breiman, L., *Bagging predictors.* Machine Learning, 1996. **26**: p. 123-140.
228. Freund, Y. and R. Schapire, *A decision-theoretic generalization of online learning and an application to boosting.* Journal of Computer and System Sciences, 1997. **55**: p. 119-139.
229. Tibshirani, R., *Regression shrinkage and selection via the lasso. .* Journal of the Royal Statistical Society, Series B, 1996. **58**: p. 267-288.
230. Efron, B., T. Hastie, I. Johnstone, and R. Tibshirani, *Least angle regression.* Annals of Statistics 2004. **32**: p. 407-499.
231. Fan, J. and R. Li, *Variable selection via nonconcave penalized likelihood and its oracle properties. .* Journal of the American Statistical Association 2001. **86**: p. 1348-1360.
232. Zou, H., *The adaptive lasso and its oracle properties.* Journal of the American Statistical Association 2006. **101**: p. 1418-1429.
233. Candes, E. and T. Tao, *The Dantzig selector: Statistical estimation when p is much larger than n.* Annals of Statistics, 2007. **35**: p. 2392-2404.
234. Wahba, G., *Spline Models for Observational Data.* 1990, Philedelphia: SIAM.
235. Ogden, R.T., *Essential Wavelets for Statistical Applications and Data Analysis.* . 1997, Birkhauser, Boston.
236. Vidakovic, B., *Statistical Modeling by Wavelets.* 1999, Wiley, New York.
237. Daubechies, I., *Orthonormal bases of compactly supported wavelets.* Communications in Pure and Applied Mathematics, 1988. **41**: p. 909-996.
238. Turkheimer, F.E., J.A. Aston, M.C. Asselin, and R. Hinz, *Multi-resolution Bayesian regression in PET dynamic studies using wavelets.* Neuroimage, 2006. **32**(1): p. 111-21.
239. Alpert, N.M., A. Reilhac, T.C. Chio, and I. Selesnick, *Optimization of dynamic measurement of receptor kinetics by wavelet denoising.* Neuroimage, 2006. **30**(2): p. 444-51.
240. Ramsay, J.O. and B.W. Silverman, *Functional Data Analysis*. Second Edition ed. 2005, Springer, New York. .
241. Abramovich, F., A. Antoniadis, T. Sapatinas, and B. Vidakovic, *Optimal testing in functional analysis of variance models.* International Journal of Wavelets, Multiresolution and Information Processing, 2004. **2**: p. 323-349.
242. Ratcliffe, S.J., G.Z. Heller, and L. Leader, *Functional data analysis with application to periodically stimulated foetal heart rate data. II: Functional logistic regression.* Statistics in Medicine, 2002. **21**(8): p. 1115-27.
243. James, G.M., *Generalized linear models with functional predictors.* Journal of the Royal Statistical Society, Series B 2002. **64**: p. 411-432.
244. James, G.M. and B.W. Silverman, *Functional adaptive model estimation.* Journal of the American Statistical Association, 2005. **100**: p. 565-576.

245. Lingjærde, O.C. and K. Liestøl, *Generalized projection pursuit regression. .* SIAM Journal on Scientific and Statistical Computing 1998. **20**: p. 844-857.
246. Fan, J. and J. Lv, *Sure independence screening for ultrahigh dimensional feature space.* Journal of the Royal Statistical Society, Series B 2008. **70**: p. 849-911.
247. McCullagh, P. and J. Nelder, *Generalized Linear Models.* 1989, London: Chapman and Hall
248. Hosmer, D.W. and S. Lemeshow, *Applied Logistic Regression* Second Edition ed. 2000, Wiley, New York. .
249. Park, M.Y. and T. Hastie, *L1-regularization path algorithm for generalized linear models.* Journal of the Royal Statistical Society, Series B 2007. **69**: p. 659-677.