



Published in final edited form as:

Contemp Clin Trials. 2010 July ; 31(4): 355–377. doi:10.1016/j.cct.2010.04.004.

Comparative Efficacy and Durability of Continuation Phase Cognitive Therapy for Preventing Recurrent Depression: Design of a Double-Blinded, Fluoxetine- and Pill-Placebo–Controlled, Randomized Trial with 2-Year Follow-up

Robin B. Jarrett^a and Michael E. Thase^b

Robin B. Jarrett: Robin.Jarrett@UTSouthwestern.edu; Michael E. Thase: Thase@Mail.Med.UPenn.edu

^a Department of Psychiatry, The University of Texas Southwestern Medical Center at Dallas, 5323 Harry Hines Boulevard, Dallas, TX, 75390-9149, United States of America; 214-648-5345; fax 214-648-5340

^b The University of Pennsylvania School of Medicine, 3535 Market Street, Suite 670, Philadelphia, PA, 19104, United States of America. Philadelphia, Veterans Affairs Medical Center and University of Pittsburgh Medical Center

Abstract

Background—Major depressive disorder (MDD) is highly prevalent and associated with disability and chronicity. Although cognitive therapy (CT) is an effective short-term treatment for MDD, a significant proportion of responders subsequently suffer relapses or recurrences.

Purpose—This design prospectively evaluates: 1) a method to discriminate CT-treated responders at lower versus higher risk for relapse; and 2) the subsequent durability of 8-month continuation phase therapies in randomized higher risk responders followed for an additional 24-months. The primary prediction is: after protocol treatments are stopped, higher risk patients randomly assigned to continuation phase CT (C-CT) will have a lower risk of relapse/recurrence than those randomized to fluoxetine (FLX).

Correspondence to: Robin B. Jarrett, Robin.Jarrett@UTSouthwestern.edu.

Financial Disclosure

Throughout the course of this research, Dr. Thase served as a consultant to and was a member of various advisory boards for Eli Lilly and Company and received honoraria for talks sponsored by this company. He has or has had similar relationships with the manufacturers of other medications used to treat depression. In addition to Eli Lilly and Company, during the past 2 years Dr. Thase has consulted with, served on advisory boards for, or received honoraria for talks from: AstraZeneca, Bristol-Myers Squibb Company, Forest Laboratories, GlaxoSmithKline, Janssen Pharmaceutica, MedAvante, Inc., Neuronetics, Inc., Novartis, Pfizer Pharmaceuticals, Schering-Plough (now Merck), Shire US Inc., Supernus Pharmaceuticals, Transept Pharmaceuticals, and Wyeth Pharmaceuticals (now Pfizer). During the past 2 years, he has received grant support from Eli Lilly and Company, Forest, GlaxoSmithKline, Otsuka, and Sepracor, Inc., in addition to funding from the National Institute of Mental Health. He has equity holdings for MedAvante, Inc. and has received royalties from American Psychiatric Publishing, Inc. (APPI), Guilford Publications, Herald House, and W.W. Norton & Company, Inc. One book currently promoted by the APPI specifically pertains to cognitive therapy. Dr. Thase also discloses that his spouse is an employee of Advogent (formerly Cardinal Health), which does business with several pharmaceutical companies that market medications used to treat depression.

Dr. Jarrett's medical center receives the fees from the cognitive therapy she provides to patients. Dr. Jarrett is a paid consultant to the NIMH.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Methods—Outpatients, aged 18 to 70 years, with recurrent MDD received 12–14 weeks of CT provided by 15 experienced therapists from two sites. Responders (i.e., no MDD and 17-item Hamilton Rating Scale for Depression ≤ 12) were stratified into higher and lower risk groups based on stability of remission during the last 6 weeks of CT. The lower risk group entered follow-up for 32 months; the higher risk group was randomized to 8 months of continuation phase therapy with either C-CT or clinical management plus either double-blinded FLX or pill placebo. Following the continuation phase, higher risk patients were followed by blinded evaluators for 24 months.

Results—The trial began in 2000. Enrollment is complete (N=523). The follow-up continues.

Conclusions—The trial evaluates the preventive effects and durability of acute and continuation phase treatments in the largest known sample of CT responders collected worldwide.

Keywords

relapse prevention; recurrent depression; cognitive therapy; fluoxetine; longitudinal follow-up

Introduction

Although the past 20 years have witnessed the introduction of a number of new antidepressant medications, many people with depression prefer psychotherapy over pharmacotherapy [1–4]. Acute phase cognitive therapy (CT) is the most extensively researched psychological treatment of depression [5–7] as well as the most influential contemporary model of psychotherapy [8]. Moreover, when compared to pharmacotherapy, CT may have the greatest potential to prevent depressive relapse after time-limited therapy [9]. Nevertheless, prophylaxis following CT is not universal; in some studies, relapse and recurrence rates over 50% have been reported within one year of termination of acute phase CT [9].

To reduce the risk of depressive relapse following acute phase CT, Jarrett developed a continuation phase model of cognitive therapy (C-CT) [10–11] and established its efficacy compared to historical [12] and assessment-only [13] control groups. Jarrett et al. [13] found that C-CT resulted in a significant reduction in the risk of depressive relapse over the 8 months that patients received it. The risk of relapse and recurrence over 24 months was moderated by risk factors, including early age of onset and the pattern of patients' remission of depressive symptoms during the final weeks of acute phase CT. Specifically, the subset of patients who had a later age of onset or achieved a stable remission (defined below) by the end of 12 weeks of acute phase CT could terminate with a lower subsequent risk of relapse and recurrence. In contrast, those with an earlier age of onset or showing less stable or incomplete remission at the end of acute phase CT were at increased relapse and recurrence risk - unless they received C-CT [13–15]. For those at higher risk, the effect size of C-CT was as large as that observed in placebo-controlled continuation phase therapy studies of FDA-approved antidepressants [9,16–17].

The rationale for the current 8-month, randomized trial and the linked 2-year follow-up study is to assess the risk of relapse and recurrence following shorter and longer courses of CT and to allow a full investigation of the largest series of patients treated with research-quality CT, as a solitary intervention, studied worldwide¹. The 8-month, randomized clinical trial tested the efficacy of C-CT or active pharmacotherapy (fluoxetine; FLX) plus clinical management compared to pill placebo (PBO) plus clinical management. The additional 2-year follow-up made it possible to determine the extent to which the hypothesized, prophylactic advantage of C-CT (versus fluoxetine) persisted following treatment termination [18]. Further, the findings allow comment on the extent to which the preventive effects of C-CT, while it is received, can actually “convert” the prognosis of the higher risk patients to a level of vulnerability comparable to the lower risk cohort.

Methods

The trial and longitudinal follow-up are called: *Continuation Phase Cognitive Therapy Relapse Prevention [C-CT-RP] Trial*. It was registered at ClinicalTrials.gov (NCT00118404, NCT00183664, and NCT00218764). This relapse prevention trial followed an initial 12–14 week course of acute phase CT for adult outpatients with recurrent major depressive disorder (MDD). Acute phase responders (see *Primary Outcomes*) were stratified into higher and lower risk groups based on the extent and stability of their remission (see *Stratification and Randomization*). The randomization focused only on the patients at higher risk for relapse after acute phase therapy and compared the efficacy of 10 sessions within 8 months of C-CT [10, 14] versus clinical management plus pharmacotherapy (double-blinded FLX or identically appearing PBO) in preventing relapse or recurrence over 1 year after protocol treatments were discontinued. While such follow-up lasted 2 years, the statistical power is sufficient to compare higher risk cells over only the first year post randomization (see *Sample Size Determination*). The purpose of the trial is to conduct the first adequately controlled test of the prophylactic efficacy of C-CT among patients at higher risk of relapse and recurrence. We predict that over 20 months post randomization, C-CT will be more effective than FLX in: a) preventing relapse and recurrence and b) consolidating response into remission and recovery.

Oversight of Research with Human Participants

This research was approved by Institutional Review Boards (IRBs) at each site (see below) prior to initiation and will be reviewed annually through completion. Adverse events have been and will be reported to each IRB and to the project's Data Safety and Monitoring Board (DSMB). The DSMB used no *a priori* stopping guidelines, and has voted annually to continue data collection. All potential participants who were seen in research clinics provided written HIPAA authorization and written informed consent for evaluation and treatment. Prior to study entry, participants viewed a standardized videotape concerning the risks of recurrent depression, its treatment, and the requirements of participation in this randomized trial and follow-up.

Data Collection Sites

The patients were recruited by teams in the Department of Psychiatry, Psychosocial Research and Depression Clinic at The University of Texas Southwestern Medical Center at Dallas (Principal Investigator: Robin B. Jarrett, Ph.D.) and in the Mood Disorders Treatment Research Program at the Western Psychiatric Institute and Clinic of the University of Pittsburgh Medical Center (Principal Investigators: Michael E. Thase, M.D. and Edward Friedman, M.D.). Recruitment methods included project promotion through IRB approved ads on the internet and in newspapers, churches, hospitals, clinics, and other community settings. Patients were

¹We found no larger samples of patients with recurrent MDD treated with research-quality CT, as a solitary intervention, in our review of the literature. The review included: (a) Jarrett and Rush's (1994) report of studies of short-term psychotherapy for depressive disorders between January 1, 1967 and November 30, 1993 [60] and (b) a new literature search between December 1, 1993 and April 1, 2010. This search included the terms "short-term psychotherapy," "acute phase," "cognitive therapy," "behavior therapy," and "interpersonal psychotherapy" in conjunction with "depression" to search for studies of patients treated with research-quality CT in PubMed, PsychINFO, and OVID. Other large samples include (but are not limited to) the ENRICHD randomized controlled trial (RCT) of 2,481 patients who suffered myocardial infarction, of which 498 patients with major or minor depression received CBT alone and 427 patients with major or minor depression and low perceived social support (LPSS) received cognitive behavioral therapy (CBT) and LPSS counseling [61]; the Bristol-Myers RCT of patients with chronic major depressive disorder (MDD) (included a total of 228 patients who received the Cognitive Behavioral-Analysis System of Psychotherapy (CBASP) alone and 227 patients who received CBASP plus nefazodone) [62]; The DELTA group who studied recurrent MDD in 172 patients, 88 of whom received preventive CBT [63]; Craigie and Nathan (2009) who compared individual CBT to group CBT in 234 outpatients with primary mood or anxiety disorders [64]; DeRubeis et al. (2005) who reported on 240 patients with moderate to severe MDD, 60 of whom received CBT alone [65]; and finally the STAR*D RCT of depressed patients with MDD who were refractory to antidepressant medications included 36 patients who received CBT alone and 65 patients who received citalopram plus CBT [66].

both self- and practitioner-referred. Patient recruitment lasted from January 3, 2000 through July 30, 2008.

Site Coordination

The Principal Investigators (Jarrett & Thase) began meeting in late 1996 to design the study. All three Principal Investigators will continue to meet by telephone regularly (i.e., nearly weekly) until study completion. Similarly, research coordinators and data managers meet by telephone routinely. The sites share data through a protected website and share information through e-mail communication between telephone meetings. Each research procedure has been written in detail to help personnel implement the procedures in a consistent manner across sites.

Inclusion/Exclusion Criteria

Consenting male or female outpatients with recurrent MDD, as diagnosed by the Structured Clinical Interview for DSM-IV (SCID-I) [19], were included when they: a) remitted between depressive episodes, had at least one prior episode with complete inter-episode recovery, or had antecedent dysthymic disorder; and b) scored 14 or more on the 17-item HRSD at the initial and second interview. Excluded patients: (a) had severe or poorly controlled concurrent medical disorders that may cause depression or require medication that could cause depressive symptoms; (b) suffered from the following concurrent DSM-IV psychiatric disorders: any psychotic or organic mental disorder, bipolar disorder, active alcohol or drug dependence, primary (i.e., predominant) obsessive compulsive disorder or eating disorders; (c) could not complete questionnaires written in English; (d) were an active suicide risk; (e) had previously not responded to a trial of at least 8 weeks of CT conducted by a certified therapist; (f) had previously not responded to at least 6 weeks of 40 mg of FLX; (g) were pregnant or planned to become pregnant during the first 11 months after intake; or (h) did not provide informed consent. Excluded patients, including those with active suicide risk, were referred for non-protocol treatment (e.g., hospitalization and/or pharmacotherapy).

Medical Evaluation and Drug Washout

When a review of medical history necessitated doing so, a physical examination and appropriate laboratory tests were obtained to ensure that patients were diagnostically eligible. Patients who were taking psychotropic medication were withdrawn from the medication and had to be unmedicated for at least one week prior to study entry.

Study Design

Figure 1 illustrates the randomized, parallel group design in which patients in the higher risk stratum were randomized to one of three arms (1:1:1) for 8 months of continuation phase therapy. Figure 1 also shows that the lower risk stratum was assessed at the same time points but was not part of the randomized design. This study was powered to test the following *Primary Hypothesis* concerning the higher risk stratum: compared to FLX plus clinical management, C-CT will significantly reduce relapse and recurrence rates in higher risk patients from randomization through one year after continuation phase protocol treatments are stopped (i.e., over 20 months). Secondary hypotheses include: *Secondary Hypothesis A*: patients who receive either of the *active* continuation phase treatments (i.e., C-CT or FLX plus clinical management) will have a significantly lower risk of relapse over the 8 months than patients randomized to the PBO plus clinical management arm; *Secondary Hypothesis B*: over the 8 months that C-CT is in effect, it will “convert” the relapse risk of higher risk patients to a rate comparable to that of the lower risk patients for the same 8 month interval; *Secondary Hypothesis C* regarding the unrandomized lower risk stratum: the risk of relapse and recurrence

over the full 32-month observational period will be significantly less for the lower risk group than for the higher risk comparators.

Sample Size Determination

The sample size was set to test the study's primary hypothesis. The hypothesized rates of relapse/recurrence were 30% for C-CT and 60% for FLX one year after discontinuation of therapy, based on previous studies [13,20]. A minimum sample size of 180 randomized patients was necessary to detect a significant difference in a one-sided log rank test where alpha was 0.05 and power was 80%. In order to accommodate the original estimates that 80% of the responders would fall into the higher risk stratum and 35% would drop-out during the experimental phase, we estimated that no less than 225 responders would need to consent to observation. We estimated that the response rate to acute phase CT would be approximately 50%, which suggested that no less than 450 patients would need to consent to participate in acute phase CT. The randomized controlled trial (RCT) was not powered to test the secondary hypothesis A.

Study Phases

As can be seen in Figure 1 and Table 1, the overall study has four phases: (a) initial evaluation, (b) approximately 12–14 weeks of CT, (c) 8 months of continuation phase treatment (higher risk patients) or follow-up only (lower risk patients), and (d) 2 additional years of follow-up free of protocol treatment. Blinded evaluations are conducted every 4 months for 32 months, with the first blinded evaluation falling at the end of acute phase cognitive therapy (see Table 1).

Acute Phase Cognitive Therapy (CT)

Acute phase CT consisted of a 16–20 individual session protocol conducted according to the treatment manual of Beck et al. [21] which was videotaped. By protocol, a maximum of 14 weeks was permitted to allow missed appointments to be rescheduled. Sessions 1 – 8 occurred twice weekly. Thereafter, patients who experienced at least a 40% reduction in HRSD-17 scores began weekly sessions and were deemed “early responders”. To maximize the likelihood of response and subsequent participation in the continuation phase trial, patients who experienced <40% reduction in HRSD-17 scores at Session 9 (compared to HRSD-17 scores from the diagnostic assessment) were deemed “late responders” and continued twice weekly sessions for four more weeks. The 40% threshold was chosen based on our clinical experience.

Therapist Training and Longitudinal Monitoring of Competence

Before therapists treated protocol patients, they were required to demonstrate proficiency in (a) CT and C-CT, as defined by the site supervisors' judgment and by maintaining Cognitive Therapy Scale (CTS) [22] scores above 39 over time; and (b) their clinician ratings for the HRSD-17 and DSM-IV diagnoses of MDD. Experienced faculty led the weekly group supervision sessions at each site and provided, when needed, individual case consultation. Videotaped sessions were selected at random for review in supervisory groups. The CTS was completed by therapy supervisors and their teams to provide feedback to therapists and to assess their competence. Dr. Jarrett attended supervision at the Dallas site and was available to therapists at both sites as needed. Sander Kornblith, Ph.D. served as the on-site CT supervisor for Pittsburgh. The Principal Investigators and Data Safety and Monitoring Board (DSMB) monitored the quality of the CTS scores across sites and time.

Cognitive Therapists' Characteristics

Fifteen therapists participated, including 11 females and 4 males. At the Dallas site all therapists were PhDs; at the Pittsburgh site, three were PhDs, one was an MD, and four were master's

level. Patients were assigned to therapists based on geographic convenience and therapist availability.

Psychoeducational Visits

Patients attended psychoeducational sessions before CT began and again within approximately 7 days of completing CT Session 11 (i.e., about Week 7 or 8 for early responders; about Week 6 or 7 for late responders). The purpose of these visits was to: (a) provide factual information about the risk of relapse/recurrence in recurrent MDD, (b) review the “road map” of treatment provided in the study, (c) verify continued consent for participation, and (d) collect self-report questionnaires.

Stratification and Randomization

On the basis of the final seven acute phase CT HRSD-17 evaluations (including the first blinded evaluation), responders were stratified into lower risk and higher risk groups. Within each site, the higher risk stratum was defined by any of the final seven consecutive HRSD-17 scores of ≥ 7 ; the lower risk group had all of the final seven HRSD-17 scores of < 7 . The higher risk patients were randomized to either C-CT or blinded FLX or PBO plus clinical management (see below). Previous findings by our teams had shown that patients with this trajectory of response were approximately four times more likely to relapse (e.g., 40% risk vs. 10% risk) following termination of acute phase CT than were patients with rapid, complete, and sustained remissions [13,23]. At the randomization, treatment assignment was also stratified by site, by each patient’s number of prior depressive episodes (i.e., 1 or 2 versus ≥ 3), and by the presence versus absence of antecedent DSM-IV dysthymia.

A study statistician oversaw the design of the stratification and blocking, and used a computer-generated random number sequence to assign patients to cells. The randomization plan was housed at the Pittsburgh site and was only available to the pharmacists at each site. The research coordinators telephoned the pharmacy to obtain instructions to randomize the patient to either C-CT or “medication clinic” (i.e., clinical management plus FLX or PBO). The pharmacists could break the blind (i.e., assignment to FLX or PBO) to non-protocol staff if a medical emergency occurred. Except for the dispensing pharmacists, all research personnel have been kept blinded to assignment to FLX or PBO.

Continuation Phase Methods for the Higher Risk Stratum

Continuation Phase Cognitive Therapy (C-CT)—With few exceptions (e.g., therapist maternity leave), the same therapist conducted the acute and continuation phases of CT. Continuation phase CT [10,14] consisted of 10, 60-minute sessions spread over 8 months (i.e., every other week for the first 2 months, followed by six monthly sessions). The primary aim of C-CT was to prevent relapse and to promote full remission and recovery. Patients were taught to use symptoms and emotional distress to activate their use of coping skills learned in acute phase CT. Symptom-reducing cognitive-behavioral techniques (e.g., cognitive restructuring and behavioral activation) were used; the more symptomatic the patient, the more the therapist structured the session. The therapists and patients collaborated to reduce residual symptoms, improve coping with adversity, decrease the probability of stressful events, and enhance behavioral and cognitive strengths. The therapists helped the patients identify and master the cognitive behavioral strategies that they found most effective in the acute phase for reducing distress and symptoms. They identified cognitive, interpersonal, or emotional vulnerabilities and continued to practice these self-help skills over time, applying them to new challenges as they arose or were anticipated. The patient assumed more responsibility for the session over the course of C-CT. The patient and therapist collaborated in identifying and enhancing patient strengths. The therapist concentrated on promoting generalization and maintenance of (a) compensatory skills that the patient had previously found helpful in

symptom reduction and (b) coping strategies to prepare for identified or anticipated vulnerabilities.

Clinical Management Plus FLX or PBO—After an initial 30–45 minute session, patients were seen for 15–30 minute sessions at the same frequency as C-CT. Sessions were conducted by experienced clinicians according to a revised version of the manual by Fawcett et al. [24] Clinicians were not permitted to use any C-CT intervention and instead focused on supportive contact around the signs and symptoms of the illness, beneficial and untoward effects of the medication, and information about depression. Side effect ratings were collected at each session using the Somatic Symptom Scale [25] and scored using the UKU (Udvalg for Kliniske Undersøgelser) method [26]. Fluoxetine was chosen because of established continuation phase efficacy [20] and a low risk of discontinuation syndromes [27]. Research pharmacies at both sites dispensed medication in unit-dose bottles. Patients initially received either FLX 10 mg/day or identically appearing PBO capsules. When necessary, the initial FLX dose could be decreased to every other day to lessen side effects. When there were no dose-limiting side effects, doses of 20 mg and 40 mg/day were prescribed at the second and third visit, respectively. These dosages were chosen on the basis of prior research in which FLX at 40 mg/day offered a balance of efficacy and tolerability [20,28]. Patients who could not tolerate this dosage remained in the study on the maximally tolerated dose (minimum dose: 10 mg/day). Patients who could take any dosage of FLX or PBO were followed for supportive care by treating clinicians. During the first six years of the project, fluoxetine and identically appearing placebo capsules were provided by Eli Lilly and Company at no cost to the project. Thereafter, fluoxetine and identically appearing placebo capsules were purchased and formulated by the research pharmacy at the Dallas site.

Quality Control of Pharmacotherapy

The integrity of the pharmacotherapy condition was ensured by using experienced clinicians and a standardized supportive clinical management protocol [24]; Dr. Thase was available for consultation as needed. In addition, medication adherence was monitored at each visit by the attending clinician and research pharmacy (i.e., pill counts).

Pharmacotherapist Characteristics

Seven male psychiatrists (two in Dallas and five in Pittsburgh) prescribed double-blinded FLX or PBO.

Lower Risk Stratum

Lower risk patients were followed for 32 months after the first blinded evaluation (i.e., beginning at the date of randomization). They were seen by a blinded evaluator at four month intervals (i.e., Months 4, 8, 12, 16, 20, 24, 28, and 32 or exit). Throughout follow-up, patients in the lower risk stratum agreed to refrain from taking mood altering medications or drugs; the evaluator asked about such consumption at each visit.

Blinded Clinical Evaluation

A clinical evaluator (masked to risk and treatment assignment) collected symptom severity measures and assessed diagnostic status according to DSM-IV criteria for a major depressive episode using the SCID-I [19] and the Psychiatric Status Ratings (PSR) within the Longitudinal Interval Follow-up Evaluation (LIFE-II) [29]. These videotaped evaluations were conducted: (a) at the end of the acute phase (i.e., within approximately 7 days of their last acute treatment session), (b) any time the patient exited the protocol, (c) at Months 4 and 8 of the continuation phase, (d) at Months 12, 16, 20, 24, 28, and 32 during the longitudinal follow-up, and (e) at any other time a relapse or recurrence was suspected. In order to promote complete data

collection, patients were paid approximately \$20 per blinded evaluation. The following safeguards were used to preserve the blind: (a) each site had at least two evaluators who were “housed” outside of the research clinic, (b) each evaluation began by reminding patients not to disclose treatment assignments, and (c) if the blind was broken for one evaluator, the second collected the remaining assessments.

Twenty-Four Month Longitudinal Follow-Up

At the end of the continuation phase, which was 8 months post-acute phase CT, both higher risk and lower risk patients were followed for 2 years, (i.e., 32 months with the count beginning at the first blinded evaluation). In Dallas only, higher risk patients were seen at two post treatment clinic visits (monthly for two months) at the beginning of the longitudinal follow-up. Patients were encouraged to contact study staff if they suspected a relapse or recurrence, at which time a blinded evaluation occurred. When the DSM-IV criteria for a major depressive episode (i.e., MDE) were satisfied according to the blinded evaluator, protocol staff encouraged patients to begin appropriate extra protocol treatment and assisted them in locating such treatment. Blinded evaluations were conducted as described above. Throughout follow-up, and similar to lower risk patients, the higher risk patients agreed to refrain from taking mood altering medications or drugs. The evaluator asked about and recorded consumption at each visit (see *Non-Protocol Treatment*). Patients continued in follow-up after relapse and recurrence, regardless of whether they initiated non-protocol treatment.

Primary Outcomes

Relapse and recurrence were the primary dependent variables within the follow-up. The definitions of these and other dichotomous outcomes are as follows:

Response to acute phase CT is defined by an independent evaluator at the end of the acute phase as: (a) the absence of DSM-IV MDE and (b) a 17-item HRSD ≤ 12 . This relatively broad definition of response determines eligibility to enter the continuation phase study (for higher risk patients) or follow-up (for lower risk patients).

Stable Remission represents a period of time that the patient is not fully symptomatic and is the operational definition of lower risk. Stable remission is defined by: (a) the last seven consecutive HRSD-17 scores < 7 during the acute phase or (b) return to “usual self” according to the LIFE, as denoted by six Psychiatric Status Ratings (PSR) ≤ 2 over 6 weeks after randomization, according to the blinded evaluator.

Relapse denotes a continuation or exacerbation of the presenting episode. Relapse is defined by a blinded evaluator as meeting DSM-IV criteria for MDD (i.e., LIFE PSR score of 5 or 6 for 2 consecutive weeks) that occur before the criteria for recovery (see below) have been met.

Recovery, the end of an episode, is a remission lasting ≥ 8 consecutive months.

Recurrence, a new episode, is defined by a blinded evaluator as meeting DSM-IV criteria for MDD (i.e., LIFE PSR score of 5 or 6 for 2 consecutive weeks), which occurs after recovery.

Evaluator Training and Calibration

To maintain a high level of diagnostic reliability within and across the sites, evaluators (i.e., diagnostic evaluators, blinded evaluators, pharmacotherapists, and cognitive therapists) participated in formal training on the use of the SCID-I [19]. This training consisted of observing and being observed by highly reliable and experienced evaluators until the trainee achieved agreement with ratings, history of illness descriptors, and DSM-IV diagnoses.

After evaluators were trained, the sites completed inter- and intra-site reliability studies on DSM-IV Current Major Depressive Disorder (MDD) diagnoses and the Hamilton Rating Scale for Depression (HRSD-17) [30]. The reliability sessions occurred regularly for the duration of the study. Videotapes were randomly selected from blinded evaluations or treatment sessions, and the evaluators on the videotapes were rotated. During each reliability session, evaluators rated the two videotapes (one from each site) in groups; a discussion of discrepant ratings was included. If an evaluator's score on the HRSD-17 differed by 4 or more points from the group mean (i.e., his/her site's mean), additional training was provided to allow his or her ratings to converge with the group's ratings, including those of the senior diagnosticians.

Intra-class correlation coefficients for the HRSD-17 and DSM-IV Current MDD diagnoses were also calculated within and across the two sites. The Principal Investigators, study teams, and the DSMB monitored the quality of the ratings across data collection.

Data Collection

Confidentiality Agreement

The National Institute of Mental Health (NIMH) issued a Certificate of Confidentiality, which protects all study participants' data until the certificate expires.

Measurement

As outlined below, 12 domains were assessed according to the frequency detailed below and in Table 2. (Note that below "Week 12" denotes the end of acute phase CT, which can occur between Weeks 12 and 14). Detailed descriptions of the measures and their psychometric properties are available in Appendix 1.

1. Demographics—Patient demographics (e.g., gender, age, sex, marital status, education) were collected at the diagnostic evaluation using the "Patient Background" self-report form.

2. Psychiatric Diagnosis—The Current Major Depressive Disorder (MDE) section of the Structured Clinical Interview for DSM-IV [19] was administered at the diagnostic evaluation, once during Weeks 4, 8, and 12 of the acute phase and at all blinded evaluations. Higher risk patients received the MDE section at all continuation phase sessions and at two post-treatment clinic visits (monthly for 2 months). Details of past depressive episodes and MDE subtyping for the current episode (e.g., recurrent, atypical) were recorded on the "MDE Specifiers and Past MDE" at the diagnostic evaluation. All other diagnoses were recorded on the "Initial Evaluation Summary."

3. Symptom Severity—The principal symptom severity measures include the Hamilton Rating Scale for Depression (HRSD-17) [30–31], Global Assessment of Functioning (GAF) [32], Inventory of Depressive Symptomatology - Self Report (IDS-SR) [33], and the Beck Depression Inventory (BDI) [34]. Sleep disturbance was assessed using the Pittsburgh Sleep Quality Index (PSQI) [35]. The reliability and validity of these measures are well established. Clinicians administered the HRSD-17 and GAF and patients completed the BDI and IDS-SR at diagnostic evaluation, follow-up interview, weekly during acute phase treatment, and at all blinded evaluations (i.e., Months 0, 4, 8, 12, 16, 20, 24, 28, and 32 or exit). Higher risk patients completed the BDI and IDS-SR and received the HRSD-17, GAF, and Somatic Symptoms Scale (SSS; to assess side effects) [25] at all continuation phase sessions and at two post-treatment clinic visits (monthly for 2 months). The PSQI was completed at the follow-up interview, CT Session 1, and acute phase CT Weeks 4, 8, and 12. Higher risk patients completed the PSQI at continuation phase Sessions 1, 3, 5, and 6–10 and two post treatment clinic visits. The PSQI was also completed at the blinded evaluations for Months 12–32 or exit. Higher

scores on the HRSD-17, IDS-SR, BDI, PSQI, and SSS indicate more severe depressive symptoms. The closer a GAF score is to 100, the higher the patient's level of overall functioning. The 17-item version of the HRSD-17, 30-item version of the IDS-SR, 21-item version of the BDI, 11-item version of the PSQI, and the 24-item version of the SSS were used in this protocol.

4. Cognitions—Self-report measures of cognitive content included the updated Attributional Style Questionnaire (ASQ) [36], the Dysfunctional Attitude Scale (DAS - Forms A & B) [37], the Beck Hopelessness Scale (BHS) [38], and the Self Control Schedule (SCS) [39]. Patients completed these measures at intake (diagnostic evaluation or follow-up interview), at the second psychoeducational visit, and at all blinded evaluations. Higher risk patients also completed the ASQ and DAS at post-randomization Months 2 and 6 during the continuation phase. Higher scores on the two ASQ subscales indicate that the respondent's attributions for negative events are more stable and more global. Higher scores on the DAS indicate more dysfunctional attitudes. Higher scores on the BHS reflect a sense of hopelessness about the future. A higher score on the SCS connotes a higher degree of learned resourcefulness. The 12-item ASQ, 40-item DAS-A/B, 20-item BHS, and the 36-item SCS were used in this protocol.

5. Mood Induction Procedure—Consenting patients participated in a standardized challenge test of cognitive-affective reactivity [40–42] performed at the end of the acute phase, at the end of the continuation phase (Month 8), and at Month 32 or exit. Responders to acute phase CT first completed a mood Visual Analogue Scale (VAS) [43] and either Form A or B of the DAS. The VAS used here measures 100 mm from endpoint to endpoint. The descriptors 'sad' and 'happy' are located on the left and right sides respectively, with arrows indicating increasing strength of mood associated with greater distance from the center. Participants were instructed to mark an X on the line corresponding to their current mood state. Participants next listened to a piece of music (Prokofiev's orchestral introduction to the film *Alexander Nevsky*, entitled "Russia Under the Mongolian Yoke," re-mastered at half-speed), while they recalled a sad event or memory from their lives. This type of induction, combining elements of music associated with sad mood and autobiographical recall, has been found to be effective in bringing on transient dysphoric mood states [41–42,44]. After listening to the music, participants completed a second VAS and the alternate form of the DAS. After responding to the final DAS, participants rated a third VAS and were debriefed by the evaluator. The order in which DAS-A and B were presented was randomly assigned. Patients received an additional \$10 compensation for each mood induction procedure.

6. Interpersonal Functioning—The Social Adjustment Scale-Self Report (SAS-SR) [45], Dyadic Adjustment Scale [46] (DYS; completed only by patients in a committed relationship), and Inventory of Interpersonal Problems (IIP) [47] were completed at intake (diagnostic evaluation or follow-up interview), at Week 1 of acute phase CT, at the second psychoeducational visit, and at all blinded evaluations; higher risk patients also completed these measures at Months 2 and 6 (Sessions 4 & 8) during the continuation phase. A higher score on the SAS-SR represents less social adjustment. A higher score on the DYS reflects positive dyadic functioning from the perspective of one partner in a committed dyad. Higher scores on the IIP reflect greater interpersonal problems. All patients with significant others also completed the Perceived Criticism Scale (PCS) [48] at the follow-up interview, the second psychoeducational visit (approximately Week 7 of acute phase CT), and at all blinded evaluations. A higher score on the PCS reflects a higher degree of perceived criticism. The Psychosocial Functioning component of the LIFE was administered at the end of the acute phase and every 4 months thereafter. Therapeutic alliance was measured with the Working Alliance Inventory (WAI) [49], using both Client and Therapist Versions. The Therapist

Version was completed at Weeks 7 and 12 of the acute phase and, for higher risk patients in C-CT, Months 4 and 8 of the continuation phase. The Client version was completed at the second psychoeducational visit and at the end of the acute phase and, for higher risk patients in C-CT, Months 4 and 8 of the continuation phase. Higher scores on these measures reflect a stronger alliance between patient and clinician. The 56-item SAS-SR, 32-item DYS, 127-item IIP, 2-item PCS, and 36-item WAIT/C were used in this protocol.

7. Personality—The Schedule for Nonadaptive and Adaptive Personality – 2nd Edition (SNAP-2) [50] is a 390-item true-false self-report inventory developed to assess trait personality dimensions using a combination of content-based and psychometric (primarily factor-analytic) techniques. An additional 50 questions were developed by Clark to form DSM-IV diagnoses. The SNAP has a basic 3-factor structure consisting of three higher order temperament scales and 12 primary trait scales: Negative Temperament (Mistrust, Manipulativeness, Aggression, Self-Harm, Eccentric Perceptions, and Dependency), Positive Temperament (Exhibitionism, Entitlement vs. Detachment), and Disinhibition (Impulsivity vs. Propriety and Workaholism). The SNAP was completed at Weeks 1 and 12 (i.e., first or second and last acute phase CT session), the final continuation phase session (i.e., Month 8), and the final blinded evaluation (i.e., Month 32).² The SNAP also contains a TRIN scale; higher scores indicate that the patient is acquiescent or yea-saying, while lower scores indicate that the patient is more denying. The TRIN midpoint is used to determine the Invalidity Index to measure inconsistency in responses. Higher scores indicate more inconsistencies.

8. Quality of Life—The 93-item Quality of Life Enjoyment and Satisfaction Questionnaire (QOLESQ) [51] was collected at intake and at all blinded evaluations. This self-report measure successfully uncouples assessment of quality of life from syndromal level depressive symptoms. Higher level of enjoyment and satisfaction with life are reflected in higher scores.

9. Life Events—All patients completed the 12-item, self-report version of the List of Threatening Experiences (LTE) [52] at intake, at the second psychoeducational visit, and at all blinded evaluations. The greater the score on this measure, the more traumatic events have occurred.

10. Skills of Cognitive Therapy—Patients completed the Skills of Cognitive Therapy-Patient Version (SoCT-P) [53] at the second psychoeducational visit and at all blinded evaluations (i.e., Months 0, 4, 8, 12, 16, 20, 24, 28, and 32, or exit). Therapists completed the Skill of Cognitive Therapy - Observer Version (SoCT-O) [53] at Weeks 7 and 12 of the acute phase and at Months 4 and 8 (sessions 6 and 10) during C-CT.³ Using the time frame “over the last month,” the patient and therapist rated the patient’s use of compensatory skills presented during CT and rated the likelihood that the patient will relapse or recur. These scales were used to measure the extent to which the patient used cognitive therapy skills from the perspective of an observer (i.e., the therapist) and the patient his/herself, respectively. The greater the patient’s skill, the higher the score. The 35-item SoCT-P/O was used to facilitate instrument development and will subsequently be reduced to fewer items.

11. Treatment Preference—Beginning in the sixth year, the Patient Attitudes and Expectations form (PAE) [54] was obtained at the diagnostic evaluation and at Week 12 of the acute phase. Developed for use in the NIMH Treatment of Depression Collaborative Research

²Prior to April 2002, the SNAP was not collected at the beginning of the acute phase at either site and was not collected at the Month-32 blinded evaluation in Pittsburgh.

³In addition to the description above, to foster instrument development, the SoCT-P was collected in Dallas from all patients at blinded evaluations conducted during the longitudinal follow-up, when a relapse or recurrence was suspected, and at early exit. In Pittsburgh, the SoCT-P was only collected from patients randomized to C-CT at routine blinded evaluations during long-term follow-up.

Program, the PAE form was used to explore how the patient's perceptions about depression and its treatment relate to changes in the course of illness. The 46-item PAE form was collected in this protocol.

12. Non-Protocol Treatment—The Psychiatric Treatment History component of the LIFE assesses psychopharmacological and psychosocial treatment sought by the patient and was administered at the end of the acute phase and every 4 months thereafter.

Data Analysis

Overview

At randomization of the higher risk patients, treatment groups were stratified by (a) site, (b) each patient's number of prior depressive episodes (i.e., 1 or 2 versus ≥ 3), and (c) the presence versus absence of DSM-IV antecedent dysthymia. All outcome analyses will follow the intent-to-treat principle and will make use of all available data for all patients who were randomized into the continuation phase trial. We will use Cox multiple regression models to analyze times to relapse/recurrence. (Mixed models with repeated measures will be used for the HRSD-17 and other continuous variables as secondary outcomes.) Analyses of the primary hypotheses will be done using Kaplan-Meier survival curves to investigate main effects of treatment, site, and their interaction; alpha levels will be reported. The validity of the Cox proportional hazards assumption will be assessed using standard diagnostics. If there are no violations of the proportional hazards assumption, we will fit the Cox model and will formally test the primary hypothesis using this model adjusting for important confounders (e.g., site, number of previous episodes, presence of dysthymia, severity group, and therapist effects). Continuous, binary, or "count" outcomes will be analyzed using linear, logistic, or log-linear random effects models, respectively [55]. The random effects models used for these analyses will have fixed effects for group differences (e.g., non-responder vs. intermediate response; C-CT vs. pharmacotherapy) and will allow for random intercept and slopes for patients. Selected interaction effects will also be included and tested as needed. Where appropriate, different imputation strategies involving secondary outcomes and/or covariates will be compared and a sensitivity analysis will be done to assess the bias introduced by the missing data.

Model assumptions will be verified using appropriate residual diagnostic analyses. Appropriate sensitivity analyses will be performed to assess model lack-of-fit and need for alternative models.

There were no planned, unmasked, interim analyses of the primary hypothesis.

Discussion

While there is no cure for depression, cognitive therapy is the most frequently studied form of psychotherapy for depression and is particularly distinguished by its potential to have a more durable or enduring effect. Whereas the therapeutic effects of antidepressants gradually dissipate after treatment is discontinued, results of a comparative meta-analysis suggest that CT may have more durable effects [9]. However, too often prophylaxis following a time-limited course of CT is incomplete and may fade with time, leaving a significant number of CT responders at increased risk for relapse and recurrence. The field needs to better understand which parameters influence when and how long CT prophylaxis can last.

To our knowledge this study will yield the largest cohort of CT patients ever collected and will follow longitudinally the largest group of CT responders up to 2 years post-treatment. Additionally, to date this is also the only study that applies a preventive model of CT focused on relapse/recurrence [10–11] in a higher risk group of patients with recurrent MDD. Further,

this study is the first to evaluate (prospectively) the efficacy of continuation phase pharmacotherapy (FLX) after incomplete remission with acute phase CT. The results have the potential to help patients and their families by identifying effective treatment sequences that prevent relapse and recurrence and restore patients with recurrent depressive disorder to full functioning. The follow-up is important because it will allow comment not only on C-CT's preventive effect on relapse (while patients receive it) but also on recurrence (after it is discontinued).

The primary prediction is that the risks of relapse and recurrence will be lower for patients receiving C-CT than for those receiving FLX from randomization through the first year that therapies were discontinued. A secondary prediction is that relapse rates for C-CT and FLX will not differ but will be less than PBO over the 8 months of continuation phase therapy. Additionally, we also predict that the risk of relapse and recurrence (a) for the lower risk group will be significantly less than that of the higher risk group and (b) for the higher risk patients undergoing "active" C-CT (i.e., over the 8 months that patients receive it) will convert to a rate comparable to that of the lower risk patients for the comparable 8-month time period.

This study is unique in that it will establish indications for a continuation phase model of psychotherapy, designed to prevent relapse and improve longer-term outcomes. Such innovations in treatments and their selection are essential for improving the course of patients' depressive illnesses, given that many people either refuse to accept and/or fill an initial prescription or rapidly discontinue use of the antidepressant medication [56]. Such results are extremely important to the U.S. public at a time when citizens are asking pointed questions of the pharmaceutical and health care industries regarding benefit/risk appraisals, particularly in light of the FDA's warnings about the associated risk of suicidal behavior in children, adolescents, and young adults [57–58], the recent reports about the numbers of unpublished failed or negative trials of antidepressant medications, and the relationship of such publication bias to pharmaceutical industry funding [59].

Once completed, the results have the potential to: be used prospectively by therapists, lead to straightforward treatment decisions that significantly improve outcomes, yield clinically meaningful reductions in relapse and recurrence risk, and begin to understand how long preventive effects endure. Such empirically based decision rules could lower the cost of care by providing, on an individualized basis, the right amount of treatment: neither "over-treating" patients at lower risk nor "under-treating" those patients at higher risk. Such data from this ongoing, randomized clinical trial and follow-up have important implications for public health.

Acknowledgments

This report was supported by Grants Number K24-MH001571, R01 MH-58397, R01 MH-69219 (to Robin B. Jarrett, Ph.D.) and R01 MH-58356 and R01 MH-69618 (to Michael E. Thase, M.D.) from the National Institute of Mental Health. We appreciate the support of our NIMH Program Officer, Jane Pearson, Ph.D., throughout this investigation. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Mental Health (NIMH) or the National Institutes of Health (NIH). We also wish to acknowledge the unrestricted support of Eli Lilly and Company, who provided the fluoxetine and matched pill placebo until 2006. Thereafter, study materials were purchased and prepared to appear identical for both sites by the pharmacy at the University of Texas Southwestern Medical Center at Dallas.

We are grateful to our patients who made this trial possible. We appreciate the dedication and longstanding commitment of our research teams and our many colleagues (see Appendix II) at the University of Texas Southwestern Medical Center at Dallas, the University of Pittsburgh (where Dr. Thase was located during patient accrual), and the University of Pennsylvania. We also appreciate the diligence of the members of the Data Safety and Monitoring Board (see Appendix II).

References

1. Paykel ES, Hart D, Preist RG. Changes in public attitudes to depression during the Defeat Depression Campaign. *Br J Psychiatry* 1998;173:519–22. [PubMed: 9926082]
2. Crisp AH, Gelder MG, Rix S, Meltzer HI, Rowlands OJ. Stigmatization of people with mental illnesses. *Br J Psychiatry* 2000;177:4–7. [PubMed: 10945080]
3. Roper. Roper Report. New York: Roper Organization; 1986.
4. Seligman MEP. The effectiveness of psychotherapy: The consumer reports study. *Am Psychol* 1995;50:965–74. [PubMed: 8561380]
5. Clarkin JF, Pilkonis PA, Magruder KM. Psychotherapy of depression: Implications for reform of the health care system. *Arch Gen Psychiatry* 1996;53(8):717–23. [PubMed: 8694685]
6. Hollon SD, Jarrett RB, Nierenberg AA, Thase ME, Trivedi M, Rush AJ. Psychotherapy and medication in the treatment of adult and geriatric depression: Which monotherapy or combined treatment? *J Clin Psychiatry* 2005;66(4):455–68. [PubMed: 15816788]
7. Jarrett, RB.; Vittengl, JR. Effectiveness of psychological treatments for depression. In: Wells, A.; Fisher, P., editors. *Treating depression: Principles and practice of CBT, MCT and Third Wave Therapies*. Chichester, UK: John Wiley & Sons, Inc; in press
8. Norcross, JC.; Goldfried, MR. *Handbook of psychotherapy integration*. New York, NY: Basic Books; 1992.
9. Vittengl JR, Clark LA, Dunn TW, Jarrett RB. Reducing relapse and recurrence in unipolar depression: A comparative meta-analysis of cognitive-behavioral therapy's effects. *J Consult Clin Psychol* 2007;75(3):475–88. [PubMed: 17563164]
10. Jarrett, RB. Unpublished treatment manuals. Vol. 1989. 1992. *Cognitive therapy for recurrent unipolar major depressive disorder. The continuation/maintenance phase*.
11. Jarrett, RB.; Vittengl, JR.; Clark, LA. Preventing recurrent depression. In: Whisman, MA., editor. *Adapting cognitive therapy for depression: Managing complexity and comorbidity*. New York: Guilford Publications; 2008-a.
12. Jarrett RB, Basco MR, Rissler R, Ramanan J, Marwill M, Kraft D, et al. Is there a role for continuation phase cognitive therapy for depressed outpatients? *J Consult Clin Psychol* 1998;66(6):1036–40. [PubMed: 9874918]
13. Jarrett RB, Kraft D, Doyle J, Foster BM, Eaves G, Silver PC. Preventing recurrent depression using cognitive therapy with and without a continuation phase. *Arch Gen Psychiatry* 2001;58(4):381–8. [PubMed: 11296099]
14. Jarrett RB, Vittengl JR, Clark LA. How much cognitive therapy, for which patients will prevent depressive relapse? *J Affect Disord* 2008-b;111:185–92. [PubMed: 18358541]
15. Vittengl, JR.; Clark, LA.; Jarrett, RB. *Behav Res Ther*. Vol. 48. 2010. Moderators of continuation phase cognitive therapy's effects on relapse, recurrence, remission, and recovery from depression; p. 449–58.
16. Geddes JR, Carney SM, Davies C, Furukawa TA, Kupfer DJ, Frank E, et al. Relapse prevention with antidepressant drug treatment in depressive disorders: A systematic review. *Lancet* 2003;361(9358):653–61. [PubMed: 12606176]
17. Thase ME. Preventing relapse and recurrence of depression: A brief review of therapeutic options. *CNS Spectr* 2006;11(12 Suppl 15):12–21. [PubMed: 17146414]
18. Paykel ES, Scott J, Cornwall PL, Abbott R, Crane C, Pope M, et al. Duration of relapse prevention after cognitive therapy in residual depression: Follow-up of controlled trial. *Psychol Med* 2005;35(1):59–68. [PubMed: 15842029]
19. First, MB.; Spitzer, RL.; Gibbon, M.; Williams, JB. *Structured Clinical Interview for DSM-IV Axis I Disorders-Patient Edition (SCID-I/P, Version 2.0)*. New York: New York State Psychiatric Institute, Biometrics Research Department; 1996.
20. Montgomery SA, Dufour H, Brion S, Gailledreau J, Laqueille X, Ferrey G, et al. The prophylactic efficacy of fluoxetine in unipolar depression. *Br J Psychiatry Suppl* 1988;3:69–76. [PubMed: 3150694]
21. Beck, AT.; Rush, AJ.; Shaw, BF.; Emery, G. *Cognitive therapy of depression*. New York: Guilford Press; 1979.

22. Vallis T, Shaw BF, Dobson KD. The Cognitive Therapy Scale: Psychometric properties. *J Consult Clin Psychol* 1986;54(3):381–5. [PubMed: 3722567]
23. Thase ME, Simons AD, McGeary J, Cahalane JF, Hughes C, Harden T, et al. Relapse after cognitive behavior therapy of depression: Potential implications for longer courses of treatment. *Am J Psychiatry* 1992;149(8):1046–52. [PubMed: 1636804]
24. Fawcett J, Epstein P, Fiester S, Elkin I, Autry J. Clinical management-imipramine/placebo administration manual. *Psychopharmacol Bull* 1987;23(2):309–24. [PubMed: 3303100]
25. Kupfer DJ, Frank E, Thase ME. Somatic Symptom Scale. Unpublished measurement.
26. Lingjaerde O, Ahlfors UG, Bech P, Dencker SJ, Elgen K. The UKU side effect rating scale. *Acta Psychiatr Scand* 1987;334:1–100.
27. Coupland NJ, Bell CJ, Potokar JP. Serotonin reuptake inhibitor withdrawal. *J Clin Psychopharmacol* 1996;16(5):356–62. [PubMed: 8889907]
28. Wernicke JF, Dunlop SR, Dornseif BE. Low-dose fluoxetine therapy for depression. *Psychopharmacol Bull* 1988;24:183–8. [PubMed: 3290940]
29. Keller MB, Lavori PW, Friedman B, Nielson E, Endicott J, McDonald-Scott P, et al. The Longitudinal Interval Follow-up Evaluation: A comprehensive method for assessing outcome in prospective longitudinal studies. *Arch Gen Psychiatry* 1987;44(6):540–8. [PubMed: 3579500]
30. Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatry* 1960;23:56–62. [PubMed: 14399272]
31. Williams JBW. A structured interview guide for the Hamilton Depression Rating Scale. *Arch Gen Psychiatry* 1988;45(8):742–7. [PubMed: 3395203]
32. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4. Washington, DC: American Psychiatric Association; 1994.
33. Rush AJ, Giles DE, Schlessler MA, Fulton CL, Weissenburger JE, Burns CT. The Inventory for Depressive Symptomatology (IDS): Preliminary findings. *Psychiatry Res* 1986;18:65–87. [PubMed: 3737788]
34. Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J. An inventory for measuring depression. *Arch Gen Psychiatry* 1961;4:561–71. [PubMed: 13688369]
35. Buysse DJ, Reynolds CFI, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: A new instrument for psychiatric practice and research. *Psychiatry Res* 1989;28:193–213. [PubMed: 2748771]
36. Dykema J, Bergbower K, Doctora JD, Peterson C. An Attributional Style Questionnaire for general use. *J Psychoeduc Assess* 1996;14(2):100–8.
37. Weissman AN. The Dysfunctional Attitudes Scale: A validation study (Doctoral dissertation, University of Pennsylvania, 1979). *Dissertation Abstracts International* 1979;40:1389B–90B.
38. Beck AT, Weissman A, Lester D, Trexler L. The measurement of pessimism: The Hopelessness Scale. *J Consult Clin Psychol* 1974;42:861–5. [PubMed: 4436473]
39. Rosenbaum M. A schedule for assessing self-control behaviors: Preliminary findings. *Behav Ther* 1980;11:109–21.
40. Clark DM, Teasdale JD. Constraints of the effects of mood on memory. *J Pers Soc Psychol* 1985;28:1595–608.
41. Segal ZV, Gemar Z, Williams S. Differential cognitive response to a mood challenge following successful cognitive therapy or pharmacotherapy for unipolar depression. *J Abnorm Psychol* 1999;108(1):3–10. [PubMed: 10066988]
42. Segal ZV, Kennedy S, Gemar M, Hood K, Pedersen R, Buis T. Cognitive reactivity to sad mood provocation and the prediction of depressive relapse. *Arch Gen Psychiatry* 2006;63(7):749–55. [PubMed: 16818864]
43. Grossberg JM, Grant BF. Clinical psychophysics: Applications of ratio scaling and signal detection methods to research on pain, fear, drugs, and medical decision making. *Psychol Bull* 1978;85(5):1154–76. [PubMed: 704721]
44. Martin M. On the induction of mood. *Clin Psychol Rev* 1990;10:669–97.
45. Weissman MM, Bothwell S. Assessment of social adjustment by patient self-report. *Arch Gen Psychiatry* 1976;33(9):1111–5. [PubMed: 962494]

46. Spanier GB. Measuring dyadic adjustment: New scales for assessing the quality of marriage and similar dyads. *J Marriage Fam* 1976;38:15–28.
47. Horowitz LM, Rosenberg SE, Baer BA, Ureno G, Villaseñor VS. Inventory of Interpersonal Problems: Psychometric properties and clinical applications. *J Consult Clin Psychol* 1988;56(6):885–92. [PubMed: 3204198]
48. Hooley J, Teasdale J. Predictors of relapse in unipolar depressives: Expressed emotion, marital distress, and perceived criticism. *J Abnorm Psychol* 1989;98(3):229–35. [PubMed: 2768657]
49. Horvath AO, Greenberg LS. Development and validation of the Working Alliance Inventory. *J Couns Psychol* 1989;36:223–33.
50. Clark, LA.; Simms, LJ.; Wu, KD.; Casillas, A. The Schedule for Nonadaptive and Adaptive Personality. 2. Minneapolis, MN: University of Minnesota Press; (SNAP-2) in press
51. Endicott J, Nee J, Harrison W, Blumenthal R. Quality of Life Enjoyment and Satisfaction Questionnaire: A new measure. *Psychopharmacol Bull* 1993;29(2):321–6. [PubMed: 8290681]
52. Brugha TS, Bebbington P, Tennant C, Hurry J. The List of Threatening Experiences: A subset of 12 life event categories with considerable long-term contextual threat. *Psychol Med* 1985;15:189–94. [PubMed: 3991833]
53. Jarrett, RB.; Vittengl, JR.; Thase, ME.; Clark, LA. Skills of Cognitive Therapy (SoCT): A new measure of patients' comprehension and use. In preparation
54. Sotsky SM, Glass DR, Shea TS, Pilkonis PA, Collins JF, Elkin I, et al. Patient predictors of response to psychotherapy and pharmacology: Findings in the NIMH Treatment of Depression Collaborative Research Program. *Am J Psychiatry* 1991;148:997–1008. [PubMed: 1853989]
55. Bruce ML, Ten Have TR, Reynolds CFI, Katz IR, Schulberg HC, Mulsant BH, et al. A randomized trial to reduce suicidal ideation and depressive symptoms in depressed older primary care patients: The PROSPECT Study. *JAMA* 2004;291:1081–91. [PubMed: 14996777]
56. Lin EHB, Von Korff M, Katon W, Bush T, Simon GE, Walker E, et al. The role of the primary care physician in patients' adherence to antidepressant therapy. *Med Care* 1995;33(1):67–74. [PubMed: 7823648]
57. Brent D. Antidepressants and suicidal behavior: Cause or cure? *Am J Psychiatry* 2007;164(7):989–91. [PubMed: 17606644]
58. Posner K, Oquendo MA, Gould M, Stanley B, Davies M. 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):1035–43. [PubMed: 17606655]
59. Turner EH, Matthews AM, Linardatos E, Tell RA, Rosenthal R. Selective publication of antidepressant trials and its influence on apparent efficacy. *N Engl J Med* 2008;358(3):252–60. [PubMed: 18199864]
60. Jarrett RB, Rush A. Short-term psychotherapy of depressive disorders: Current status and future directions. *Psychiatry: Interpersonal and Biological Processes* 1994 May;57(2):115–32.
61. Berkman LF, Blumenthal J, Burg M, Carney RM, Cowan MJ, Czajkowski SM, et al. Effects of treating depression and low perceived social support on clinical events after myocardial infarction: The Enhancing Recovery in Coronary Heart Disease Patients (ENRICH) randomized trial. *JAMA* 2003;289(23):3106–16. [PubMed: 12813116]
62. Keller MB, McCullough JP, Klein DN, Arnow B, Dunner DL, Gelenberg AJ, et al. A comparison of nefazodone, the cognitive behavioral-analysis system of psychotherapy, and their combination for the treatment of chronic depression. *N Engl J Med* 2000;342(20):1462–70. [PubMed: 10816183]
63. Bockting CL, Spinhoven P, Koeter MW, Wouters LF, Schene AH. Prediction of recurrence in recurrent depression and the influence of consecutive episodes on vulnerability for depression: A 2-year prospective study. *J Clin Psychiatry* 2006;67(5):747–55. [PubMed: 16841624]
64. Craigie MA, Nathan P. A nonrandomized effectiveness comparison of broad-spectrum group CBT to individual CBT for depressed outpatients in a community mental health setting. *Behav Ther* 2009;40:302–14. [PubMed: 19647531]
65. DeRubeis RJ, Hollon SD, Amsterdam JD, Shelton RC, Young PR, Salomon RM, et al. Cognitive therapy vs medications in the treatment of moderate to severe depression. *Arch Gen Psychiatry* 2005;62:409–16. [PubMed: 15809408]

66. Thase ME, Friedman ES, Biggs MM, Wisniewski SR, Trivedi MH, Luther JF, et al. Cognitive therapy versus medication in augmentation and switch strategies as second-step treatments: A STAR*D report. *Am J Psychiatry* 2007;164(5):739–52. [PubMed: 17475733]

APPENDIX I

Description of Instruments

A. Demographics

1. Patient Background—The patient uses this form to report demographics (i.e., age, gender, ethnicity, marital status, etc.), use of medications, drugs, and alcohol, medical history, and reproductive status.

B. Psychiatric Diagnosis

1. Structured Clinical Interview for DSM-IV with Psychotic Screen (SCID-I/P w/ PSY SCREEN)—The SCID-I/P w/PSY SCREEN [1] is a structured interview organized to assess Diagnostic and Statistical Manual (DSM-IV) ([7]) criteria for each disorder, including Major Depressive Disorder (MDD), in a systematic fashion. Reliability between raters using the previous version of the SCID-I was assessed with the Fleiss' modification of the kappa statistic to account for chance agreement and variation in rater pairs. The overall kappa was 0.74 ($p < 0.01$), with kappa for Major Depressive Disorder being 0.72 ($p < 0.05$) [2].

2. Longitudinal Interval Follow-up Evaluation (LIFE)—The LIFE [3] is a semi-structured interview with three sections: Psychiatric Status Ratings, Psychosocial Functioning Interview, and Treatment History. Ratings are collected retrospectively in weekly (Psychiatric Status Ratings and Treatment History) or monthly (Psychosocial Functioning) intervals. The LIFE has been used in numerous studies investigating the course of Major Depressive Disorder, including the NIMH Collaborative Study on the Psychobiology of Depression [4]. Keller et al. [5] reported excellent reliability for week-by-week Psychiatric Status Ratings of major episodic affective illnesses with interclass correlation coefficients of at least .90. Reliability for other disorders evaluated using the Research Diagnostic Criteria was in the good range with interclass correlation coefficients (ICCs) of at least .70. In general, reliability among psychosocial items was high, with quantitative items showing excellent reliability. While all items showed acceptable reliability, items measuring current functioning were higher than items measuring past functioning. Warshaw, Keller, and Stout [6] reported good to excellent agreement on minimum and maximum Psychiatric Status Ratings in the first 8 weeks or the follow-up interval and for the 8 weeks preceding the interview. Test-retest reliability for 12-month Psychosocial Functioning yielded ICCs from 0.57 to 0.81. Summary measures derived from Psychiatric Status Ratings were correlated with patients' Global Assessment of Functioning Scale (GAS) [7] and the emotional function summary scores from the Medical Outcome Survey (MOS) [8], demonstrating good external validity ($r = -.45$ to $-.57$). For Treatment History, Warshaw, Keller, and Stout [6] reported very good agreement on whether or not a medication was received, with kappas ranging from 0.85 to 1.0. Dosage also yielded good to excellent agreement, with ICC of 0.80 for the first week of the follow-up interval and 0.79 for the last week of the follow-up interval. Agreement on psychosocial treatments was good with kappas from 0.68 to 0.88, depending on treatment modality. Researchers have shown that the LIFE is a reliable instrument for use in longitudinal studies, when used with an intensive training and rater monitoring system [9].

The LIFE Psychiatric Status Ratings track the longitudinal course of DSM-IV Axis I psychiatric disorders. Major Depressive Disorder is rated on a 6-point scale according to duration and severity with higher scores indicating greater depressive symptomatology. Scores of 5 and 6 indicate that the disorder is present according to

DSM-IV criteria. Other DSM-IV diagnoses are rated on a 3-point scale with a 3 indicating the presence of a disorder.

The LIFE Psychosocial Interview assesses the longitudinal course of psychosocial functioning. Patients identify their primary roles (e.g., full-time employed, part-time employed, unemployed, homemaker) and how well they perceive themselves to be performing in the areas of work, household duties, and student activities. This measure also assesses their quality of interpersonal relationships, changes in marital status, sexual satisfaction, global satisfaction, involvement in recreational activities, overall social adjustment (as rated by the patients), and the clinician's assessment of the patient's level of social adjustment. Items assessing impairment or dissatisfaction are rated on a 5-point scale with higher scores indicating greater impairment or dissatisfaction. Most items assessing quality of interpersonal relationship are rated on a scale from 0 to 7. A score of 0 indicates that the relationship is not applicable. Scores of 6 and 7 indicate variability in the quality of relationship among relationships in the category (e.g., patient has a good relationship with one sibling and a poor relationship with another). Scores from 1 to 5 indicate a continuum of quality of relationship with higher scores representing poorer quality.

The LIFE Psychiatric Treatment History assesses psychopharmacological and psychosocial treatment sought by the patient. Psychopharmacological treatment is coded for the type of medication, the psychiatric or psychological diagnosis or problem for which treatment was sought, and the dosage of the medication taken per day. Psychosocial treatments are recorded for the type of treatment sought (e.g., marital therapy, individual therapy, self-help), the psychiatric or psychological diagnosis or problem for which treatment was sought, the frequency of treatment, and who was the focus of treatment (e.g., patient, patient's child). Any psychiatric treatment the patients reported consuming outside of the protocol was recorded on the LIFE.

C. Symptom Severity

1. Hamilton Rating Scale for Depression (HRSD)—The HRSD [10] is a clinician rating scale designed to assess the severity of depressive symptoms in patients already diagnosed with MDD; higher scores reflect greater symptom severity. Each of the 17 items is rated by the clinician on either a 3- or 5-point scale, and the total score is determined by summing the item scores. Scores >24 indicate severe depression (typical of inpatients), while scores <17 suggest mild symptoms (more typical of outpatients), and scores <6 are considered to indicate an absence of depressive symptoms. With highly trained raters, the HRSD and similar specific depression symptom measures have been found to have a good inter-rater reliability ($r = 0.85$) [11]. Similarly, Kobak et al. [12] found that raters who were experienced and calibrated had a higher inter-rater reliability ($r = 0.93$) than those who were inexperienced or those who were experienced but not calibrated. There are few data on the internal consistency of the measure, but Schwab, Bialow and Holzer [13] found that individual items correlated with total score 0.45 to 0.78. Regarding validity, Knesevich, Biggs, Clayton, and Ziegler [14] found HRSD change scores to be correlated 0.68 with global change scores, and numerous studies have shown significant differences in the HRSD scores of normal controls and patients with depression, supporting its criterion validity. The measure has also shown good convergent validity with other clinical self-report depression measures ($r_s =$ approximately 0.83 and 0.70, respectively) [11]. The 17-item version of the HRSD was scored in this protocol.

2. Beck Depression Inventory (BDI)—Perhaps the most widely used measure of depressive symptoms is the BDI [15]. This 21-item self-report measure was designed to assess the intensity of depressive symptoms in psychiatric patients and to detect depressive symptoms

in normal populations. Items are rated on a 0- to 3-point scale with higher numbers indicating greater symptom severity. Cut-off scores are as follows: less than 10, none or minimal depression; 10–18, mild-to-moderate depression; 19–29, moderate-to-severe depression; and 30–63 indicates a severe depression. Beck, Steer and Garbin [16] summarized 25 years of research on the BDI. Across studies using various subject populations, they found an average internal consistency of 0.87 (range = 0.76 to 0.95) and an average short-term (<1 month) test-retest reliability of 0.60 (range 0.48–0.86). They also reported considerable evidence supporting the validity of the BDI. Correlations between BDI scores and clinical ratings of depression in psychiatric patients ranged from 0.55 to 0.96 ($M = 0.72$). The BDI also has good convergent validity with the HRS-D (range = 0.61 to 0.87) and with other self-report measures of depressive symptoms [16]. In 1996, the original BDI was updated as the BDI-II [17]. The original 21-item BDI was used in this protocol.

3. Inventory for Depressive Symptomatology - Self Report (IDS-SR)—The IDS [18] was designed to measure specific signs and symptoms of depression in both inpatients and outpatients. This 28-item questionnaire was developed as both a self-report measure (IDS-SR) and a clinician rating scale (IDS-C). Each item is rated on a 0- to 3-point scale with higher scores representing greater severity. Twenty-six of the 28 items contribute to the total score, which can range from 0 to 78. The instrument was designed to assess vegetative symptoms, cognitive changes, mood disturbance, endogenous symptoms, and anxiety symptoms. Mean scores for the IDS-SR were 36.5 for patients with depression, 21.8 for other diagnostic groups, and 2.1 for normal controls. These values differ significantly and support the construct validity of the measure. The internal consistency reliability of the IDS-SR and IDS-C are high with Cronbach's alpha (α) ratings of 0.85 and 0.88, respectively. The self-report version was highly correlated with the BDI ($r = 0.78$) and the HRSD ($r = 0.67$), as was the clinician rating, with correlations of 0.61 with the BDI and 0.92 with the HRSD. An advantage of the IDS is that it allows assessment of some atypical features. The Dallas group revised the IDS to include the previously omitted symptoms of leaden paralysis and rejection sensitivity, bringing the total number of questions to 30. Trivedi et al. [19] found that the IDS-SR-30 has high internal consistency at $\alpha=0.92$. They also found that item-total correlations were highest for major symptoms like sad mood, anxious mood, reactivity of mood, quality of mood, concentration/decision making, self-outlook, future outlook, involvement, energy/fatigability, and pleasure/enjoyment. The 30-item IDS-SR was used in this protocol.

4. Global Assessment of Functioning Scale (GAF)—The GAF [7] is a standardized assessment of overall functioning, included as Axis V in the DSM-IV. Scores may range from 1 to 100, with descriptive anchors at each 10-point interval. Depressed outpatients typically score between 40 and 65 before treatment, and 65 to 80 after responding to treatment. A high score on the GAF indicates good functioning, where a low score indicates poorer functioning with increased impairment. The GAF is highly correlated with its predecessor, the Global Assessment Scale [20]. Vatnaland et al. [21] found that interclass coefficients of .81 to .85 were high between researchers who rated patients. Recent research has confirmed that frequent training and rater calibration is important to increase reliability of GAF scores [21–22].

5. Pittsburgh Sleep Quality Index (PSQI)—The PSQI [23] is an 11-item self-report inventory that is reliable and sensitive to change. Each item is scored from 0 (not during the past month) to 3 (3 or more times each week). Carpenter and Androkowski [24] examined the psychometric properties of the PSQI reporting that Cronbach's alphas were consistent across 472 patients (bone marrow transplant, $n=155$; renal transplant, $n=56$; breast cancer, $n=102$; benign breast problems, $n=159$). Specifically, the global PSQI score had a Cronbach's alpha of .80, while the Cronbach's alpha for the items related to sleep disturbance ranged from .70 to .78 [24]. Symptom clusters (components) were formed by grouping individual items that

assessed a specific component of sleep quality. Correlations between global scores and symptom cluster (component) scores were moderate to high, suggesting good internal consistency. Global to symptom cluster (component) score correlations were lower (moderate) for sleep disturbance and daytime dysfunction ($r = .42-.60$). The correlations between global scores and specific cluster (component) scores, including sleep quality, sleep latency, and habitual sleep efficiency, were high, ranging between .70 and .83. A higher score indicates worse sleep quality. The 11-item PSQI was used in this protocol.

6. Somatic Symptoms Scale (SSS)—The SSS [25] is a 24-item checklist the clinician uses to record the patient's self-report (i.e., perception) of side effects from medication and/or the patient's perceived severity of somatic symptoms. The first 23 items assess somatic symptoms and include constructs such as dizziness, blurred vision, sexual difficulties, skin rash, and increased perspiration. The items are rated on a 3-point scale (0 = not present; 1 = present, but tolerable; 2 = present and causing significant distress or incapacity). Thus, higher total scores on the SSS indicate a greater number and severity of symptoms. The last item assesses drug and alcohol use since the previous visit. The 24-item SSS was used in this trial.

D. Cognitions

1. Attributional Style Questionnaire (ASQ; 1996)—The original ASQ [26] assesses the extent to which the respondent attributes the cause of six positive and six negative events to factors which are internal versus external, stable versus unstable, and global versus specific. The reliability of the three original attributional dimensions was modest, which led to the development of the Expanded Attributional Style Questionnaire (EASQ) [27]. This scale does not include any positive events and increases the number of negative events assessed from 6 to 24, which improves the reliability of the three attributional dimensions. However, the EASQ is quite lengthy, making completion time-consuming. Thus, a shortened version of the ASQ [28] was created which also preserves the improved psychometric properties of the EASQ. The new version of the ASQ covers 12 negative events, yields scores for stability and globality, and has been shown to have good reliability and validity [28]. The new shortened version of the ASQ with 12 items was used in this trial. While previous versions of the ASQ used scales of 1 to 7 for scoring, this 12-item version uses a scale of -3 to 3 to clearly illustrate the bipolarity of these factors [28]. Higher scores on the two ASQ subscales indicate that respondents' attributions are more stable and more global. The internality factor that was used on previous versions of the ASQ was deleted since it was the least reliable factor. Although the 12-item version of the ASQ has not been used frequently, internal consistencies of .80 and .76 were initially found for stability and globality factors [28]. Similar results were found by Riso et al. [29]. High scores on the 12-item ASQ have been correlated with both acute and chronic depression [30] and lower than average rehabilitation in cardiac patients [31]. Many studies claim to have used the ASQ, when in fact they used one of many various other versions. This makes it difficult to determine the reliability of the measure across different populations, etc. [32–33]. The construct validity of the ASQ has been supported in reviews [34–35].

2. Dysfunctional Attitudes Scale (DAS)—The Dysfunctional Attitudes Scale (DAS - Forms A and B) is a 40-item self-report measure of “silent assumptions” (i.e., a measure of severity of dysfunctional thoughts). Statements regarding perfectionist standards, concern about approval from others, requirements for being happy, and feelings of inadequacy are rated on a 7-point Likert scale ranging from “agree very much” to “disagree very much.” Higher scores indicate a greater number and severity of dysfunctional attitudes. Students with depressive symptoms ($BDI > 10$) had an average score of 130.26 ($SD = 29.60$), while the average score for students without depressive symptoms was 114.46 ($SD = 25.10$) [36]. The mean score for women diagnosed with depression using the SCID-I (Structured Clinical Interview for DSM-IV Axis I Disorders [SCID]; [1]) was 157.5 ($SD = 38.2$); women who never experienced

depression had a mean score of 96.9 (SD=24.2) [37]. The DAS successfully distinguishes depressed inpatients from nondepressed clinical controls [38]. High internal consistency scores have been reported (e.g., .93 to .96) [39–40]. Test-retest reliability correlations of .73 over 6 weeks [41] and .84 over 8 weeks [36] have been obtained. Both a lack of change on DAS scores during treatment and high DAS scores have been linked to shorter time before relapse [42–43]. The validity of the DAS is supported by its correlation with BDI scores (e.g., .31 to .49) [39,41,44–47]. In addition, DAS scores differ significantly between psychiatric patients with and without depression and volunteers without depression [48]. The concurrent validity of the DAS is supported by its correlation of .36 and .47 (p 's < 0.001) with two measures of depressive symptoms (D-Scale of the Profile of Mood States [POMS]; [49] and the BDI) in a student population [36]. Although the validity and reliability of the DAS is strong, there is continued debate regarding the factor structure of the scale and if it is consistent across different communities [50–51]. The DAS has been translated to Chinese, Japanese, Norwegian, and Dutch and is found to have good reliability and validity in communities using those languages, supporting the utility of the measure across cultures and in heterogeneous populations [52–54]. The 40-item DAS was used in this protocol.

3. Beck Hopelessness Scale (BHS)—The BHS was designed to measure an individual's negative expectancies about the future [55–56]. It is derived from an earlier hopelessness measure called the Generalized Expectancy Scale [57]. The scale consists of 20 true/false items. Nine of these statements are keyed false, and 11 are keyed true. A score of 0 or 1 is given to each answer; the higher the score, the greater degree of hopelessness. An individual does not possess hopelessness with a score of 3 or less and has only mild hopelessness with a score of 4–8. Scores of 9–14 points are considered moderate hopelessness, and scores above 15 points are severe hopelessness [16]. High levels of hopelessness on the BHS are significantly associated with greater severity of depression in adults [58] and children [59]. Individuals with scores in the moderate to severe range have an increased potential for suicide [60], although a review found that the degree of that risk is smaller than originally believed [61]. Interestingly, the correlation between high hopelessness scores and increased potential for suicide is not found among individuals with alcoholism [62]. A population study in Finland with 1722 respondents found that individuals are more likely to have a high hopelessness scores if they believe their financial situation, health or working ability are poor [63]. The scale is widely used and considered reliable, with the scale reporting a high internal consistency yielding a reliability coefficient (Kuder Richardson-20) of 0.93 [55]. In a non-clinical population, the internal consistency was 0.88 [64]. The 20-item BHS was used in this protocol.

4. Self-Control Schedule (SCS)—The Self-Control Schedule assesses an individual's tendencies to use self-control methods to solve behavioral problems [65]. It is considered a measure of learned resourcefulness. The 36 items on the SCS assess four categories of self-control skills: the use of cognitive strategies, the use of problem-solving strategies, the ability to delay gratification, and a general belief in one's ability to regulate internal events. Each of the items is rated on a 6-point Likert scale ranging from +3 (very characteristic of me) to –3 (very uncharacteristic of me). A higher score on the SCS connotes a higher degree of self-control. A four-week test-retest correlation ($r=.86$) indicated that the measure is fairly stable [65]. Satisfactory internal consistency scores ranging from .78 to .86 have repeatedly been reported [65–69]. While SCS scores alone have not always been predictive of response to cognitive therapy [70], patients with high scores combined with more severe depressive symptoms did respond to cognitive therapy [71]. Additionally, high SCS scores have been correlated with confidence in students [66] and with positive cognitions in elders [72]. High scores on the SCS have also been negatively correlated with scores on the BDI [69]. The SCS has been translated to Chinese and Thai and has been found to have good reliability and validity

in communities using those languages, supporting the utility of the measure [73–74]. The 36-item SCS was used in this protocol.

E. Mood Induction Procedure

1. Dysfunctional Attitudes Scale (DAS)—Described above, the DAS was used in this protocol to assess dysfunctional attitudes before and after a mood induction procedure.

2. Visual Analogue Scale (VAS)—The Visual Analogue Scale (VAS) is a Likert-type scale used to assess mood. Respondents designate their current mood on a scale measuring 100 mm end to end. While the original VAS was 156 mm end to end [43,75], many investigations, including the current protocol, have used a 100 mm long scale [76–79]. It is important to note that available psychometrics involve both lengths of scales. The descriptor ‘sad’ is located to the left of the center, and ‘happy’ is located to the right. Arrows indicate that increasing strength of mood is associated with a greater distance from the center [80]. Patients are asked to mark their current mood with an “X” on the line. The mean mood change for formerly depressed patients who completed the VAS pre- and post-mood induction at a length of 152 mm has ranged from –24.6 to –25.48 [43,75]. Williams et al. [81] found that self-described end anchors may increase the reliability of any type of linear scale, since different people may have different definitions of fixed end anchors. Results, however, suggest that fixed anchors (like the ones used in this study) are just as effective and easier to use [82–83].

F. Interpersonal Functioning

1. Dyadic Adjustment Scale (DYS)—Global relationship satisfaction was measured by the DYS [84]. The DYS is a 32-item self-report measure assessing both the quality of marriages or similar dyads and satisfaction with that relationship. The DYS is likely the most widely used measure of dyadic adjustment in research [85]. Scores on the scale ranged from 0 to 151 with mean scores of 114.8 ($SD = 17.8$) for married couples, 70.7 ($SD = 23.8$) for divorced couples, and 101.5 ($SD = 28.3$) for the total sample [84]. Higher scores reflect greater dyadic adjustment. A score of less than 97 is commonly used as a criterion of dyadic discord (i.e., one standard deviation below the mean scores of married couples) [84]. Total DYS scores have regularly differentiated distressed and non-distressed couples [86–87]. Spanier reports an internal consistency reliability of 0.96 (Cronbach’s coefficient α) and a meta-analysis of studies using the DYS found a mean internal consistency reliability of .915 across a wide range of relationship types [85]. In these studies, reliability was not substantially influenced by marital status, ethnicity, sexual orientation, or gender. Spanier [84] found evidence for content, criterion-related, and construct validity of the DYS via comparisons with the Locke-Wallace Marital Adjustment Test (MAT) [88], factor analysis of the scale, and comparisons of normal and divorced couple groups. The 32-item DYS was used in this protocol.

2. Social Adjustment Scale - Self Report (SAS-SR)—The original SAS-SR [89] is a 42-item self-report measure of instrumental and expressive role performance. The items are rated on a 5-point scale, with higher scores indicating impairment. Mean internal consistency coefficient α of 0.74 and mean test-retest reliability of 0.80 across two different time periods have been reported [90]. The concurrent validity of the SAS-SR has been demonstrated by Weissman and colleagues [90] who reported significantly different scores from psychiatric patients versus community controls, as well as from acute versus recovered depressive patients [89]. In addition, the SAS-SR has been found to be significantly correlated with clinical ratings such as the HRS-D and the Raskin Depression Scale (RDS) [91], as well as self-report measures such as the Center for Epidemiological Studies - Depression Scale (CES-D) [92] and the Symptom Checklist (SCL-90) [93]. The original SAS-SR has been translated in Japanese [94] and modified for older adults, ages 63 to 87 [95]. In addition to the original version, two other versions have been developed. These include a) the Social Adjustment Scale II (often

completed by schizophrenic patients) [96–97], and b) an abbreviated version of the SAS II—the Social Adjustment Scale for the Severely Mentally Ill (SAS-SMI), used for patients so diagnosed [98]. A higher score on this measurement represents less adjustment. The 56-item SAS-SR was used in this study.

3. Inventory of Interpersonal Problems (IIP)—The IIP [99] is a 127-item self-report measure of distress associated with interpersonal difficulties. Respondents are presented a list of common problems and are asked to consider if each has been present in their relationships with significant others. They rate the level of distress associated with each problem on a 5-point Likert scale from 0 (not at all distressing) to 4 (extremely distressing); thus higher scores signify greater levels of distress. Interpersonal problems are divided into subscales indicating difficulties in being assertive, sociable, submissive, intimate, or too responsible or controlling. Mean scores for each subscale correspond to the level of distress associated with that area and can be compared to the norms derived from nonpsychiatric or psychiatric populations [99–100]. Ten-week test-retest reliability for the overall mean score was 0.98 (Horowitz et al., 1988). Internal consistency for the subscales ranged from .82 to .94, and their 10-week test-retest reliability ranged from 0.80 to 0.90 [99]. Concurrent validity of the IIP is supported by predicted associations of IIP personality categories with other assessments, such as the Symptom Checklist-90R (SCL-90-R)[101–102] and therapist evaluations of patients' personalities [103]. The IIP is also sensitive to change resulting from psychotherapy [99]. Lower scores on the general subscales and circumplex subscales are indicative of higher levels of that trait. For example, a low score on the assertive scale indicates higher levels of assertiveness. The 127-item IIP was used in this protocol.

4. Perceived Criticism Scale (PCS)—The PCS is a measure of expressed emotion (EE) that consists of a two-question, 10-point, Likert-type scale which asks the subject to rate (a) “How critical is your spouse of you?” and (b) “How critical are you of your spouse?”[104]. These authors found that ratings of the spouse's level of criticism were a significant predictor of relapse in depressed patients [104]. The scale also demonstrated good 3-month test-retest reliability [104]. PCS patient ratings were significantly correlated with total EE as assessed by the Camberwell Family Interview [105], although not with the criticism subscale [104]. A rating of high perceived criticism does not appear to be a proxy measure of symptom severity as scores are not correlated with scores on the Beck Depression Inventory (BDI) [15], Present State Exam [104,106], Hamilton Rating Scale for Depression (HRSD) [10,107], Global Assessment of Functioning (GAF), Symptom Checklist 90-Revised (SCL-90-R) [108–109], nor the Inventory to Diagnose Depression (IDD) [110]. A review of EE assessments recommended the PCS where a brief measure of EE is needed [111]. A higher score on this scale reflects a higher degree of criticism. The two-item PCS was used in this study.

5. Working Alliance Inventory (WAI)—The WAI [112] is a 36-item self-report measure designed to assess the quality of the alliance between a therapist and patient. There is a client version and a therapist version of the WAI. The model of alliance assessed is based on Bordin's [113] tripartite conceptualization of alliance: bonds, goals, and tasks. The patient or therapist is asked to rate each of the items on a 7-point Likert scale. Each questionnaire yields a composite working alliance score and task, goal, and bond subscale scores. For the two versions of the WAI, Cronbach's alpha has ranged from 0.87 to 0.93, and the WAI has good convergent, concurrent, and predictive validity [112]. Interestingly, scores on the client version mediated the relationship between interpersonal functioning and outcome from treatment with cognitive behavioral therapy of depression [114]. Results from a metaanalysis suggested that the WAI has adequate reliability and predicts outcome. The results also lead to their recommendation of the WAI as the choice measure of therapeutic alliance [115]. A higher score on this measure

reflects a higher level of alliance between patient and clinician. The 36-item WAI was used in this protocol.

G. Personality and Personality Disorders

1. The Schedule for Nonadaptive and Adaptive Personality—2nd edition

(SNAP-2)—The SNAP-2 is a 390-item factor-analytically derived self-report inventory, with a true-false format, that assesses (a) 15 traits relevant to personality disorders (PDs), (b) the 10 *DSM-IV* personality disorders, and (c) six potentially invalidating response styles (i.e., validity scales) [116–117]. The original SNAP item pool targeted clusters of PD criteria, whereas later developments were guided by psychometric as well as content considerations. The scales form three higher order factors—Negative Temperament, Positive Temperament, and Disinhibition (vs. Constraint)—that have been identified repeatedly in factor analyses of self-report scales (e.g. [118–120]). In terms of the well-known Five-Factor Model (FFM) of personality, the SNAP factor structure is at a higher level of the personality hierarchy: Negative and Positive Temperament correspond to Neuroticism and Extraversion, respectively, whereas Disinhibition represents a combination of (low) Agreeableness and (low) Conscientiousness. Openness is minimally represented in the SNAP because it has not been found to play a major role in PD. Each factor contains a “core temperament” scale plus several more specific trait scales. Specifically, the higher order Negative Temperament factor is comprised of the core Negative Temperament scale, plus Mistrust, Manipulativeness, Aggression, Eccentric Perceptions, Self-harm, and Dependency. The Positive Temperament factor is comprised of the core Positive Temperament scale, plus Exhibitionism, Entitlement, and (on the low end) Detachment. The Disinhibition factor is comprised of the core Disinhibition scale plus Impulsivity, with Propriety and Workaholism on the other end.

Chapters 3 and 4 of the SNAP Manual [117] contain extensive psychometric data on the scales showing that they are both internally consistent (over five samples, including college student, community adults, and three patient samples; median alpha coefficient = 0.81; range = 0.63 to 0.9); [117] (Table 3.17) and have good short-term stability (Median retest $r = 0.87$ over 1 week to 4½ months; $Mdn = 7$ weeks) [117] (Table 3.18) in a community adult sample. The median pre-post retest r s were .70 and .68 in two treatment samples [117] (Table 3.19). Moreover, the 12 primary scales are much more independent than those of other instruments assessing characteristics of PDs—only 5 of 264 of the interscale correlations were stronger than $|0.50|$ in college student, community adult, or two patient samples ($Mdn = |0.20|$) in both patient and non-patient samples [117] (Tables 3.20–3.21). Conversely, supporting the construct validity of these three scales, the primary scales are moderately to strongly correlated with their respective higher order scales (Median $r = .35$, range = .13 – .70; [117] Tables 3.20–3.21).

Validity data comparing the SNAP to clinical ratings of the *DSM-IV* PDs using the Structured Clinical Interview for *DSM-IV* Personality [121] indicate good convergence between the two methods in a heterogeneous patient sample [117] (Table 4.14–4.16). For example, ratings of Narcissistic PD correlated with SNAP Exhibitionism (.50) and Entitlement (.40). Detachment correlated .56 with both Schizoid and Avoidant Personality Disorder. Mistrust correlated .52 with both Paranoid PD and Borderline PD, the latter of which also correlated with Negative Temperament (.59), Self-harm (.65), and Dependency (.50). In addition, the scales correlated systematically with such measures as Tellegen’s Multidimensional Personality Questionnaire, several measures of the five-factor model of personality, and the MMPI-2 [117]. The SNAP contains a TRIN scale. A higher score indicates that the respondent is acquiescent or yea-saying, while a lower score indicates that the patient is more denying. The TRIN midpoint is used to determine the Invalidity Index to measure inconsistency in responses. A higher score indicates more inconsistencies.

H. Quality of Life

1. Quality of Life Enjoyment and Satisfaction Questionnaire (QOLESQ)—The QOLESQ is a 93-item self-report measure designed to assess the degree of enjoyment and satisfaction experienced by subjects in multiple areas of daily functioning. Responses are grouped into eight areas of functioning which include physical health, subjective feelings, leisure time activities, social relationships, general activities, work, household duties, and school/course work. Each of the questions is scored on a 5-point scale that indicates the degree of enjoyment or satisfaction achieved during the past week. Endicott, Nee, Harrison, and Blumenthal [122] reported good test-retest reliability (range was 0.63 to 0.89 across eight subscales) and strong alpha coefficients of internal consistency (range was 0.90 to 0.96 across eight subscales). In addition, these authors found that the QOLESQ was sensitive to changes in symptom severity of depression but that it measured important dimensions of the illness that were not reflected in severity measures. Since the development of the original QOLESQ, two abbreviated versions of the measure have been developed. A 16-item form was developed from the General Activities portion of the 93-item version [122]. Another group [123] developed an 18-item abbreviated Quality of Life measure intended to identify core items from the original test which could be completed in less time. In order to increase the use of the measure, pediatric as well as Italian versions were also created [124–125]. A higher score indicates a higher level of enjoyment and satisfaction with life. The 93-item QOLESQ was used in this protocol.

I. Life Events

1. List of Threatening Experiences-Questionnaire Version (LTE-Q)—The LTE-Q [126] measures 12 categories of common life events that have been found to be threatening. Examples of categories include suffering from serious illness or injury, marital or relationship difficulties, or losing one's employment. The respondent checks off any categories of events he/she has experienced over a specified period of time prior to the assessment (e.g., 3 or 6 months). The LTE-Q provides an assessment of life events which have been identified as threatening by researchers who are unable to utilize more comprehensive interview techniques (e.g., Life Events and Difficulties Scale - LEDS, [126–127]). In fact, Brugha and colleagues [126] report that these 12 categories accounted for 77% of life events found to be related to marked or moderate long-term threats measured by more comprehensive methods [e.g. 128]. Cohen's kappa, used to measure test-retest reliability, was strong with the majority falling between 0.78 and 1.0 for each life event category [129]. The LTE-Q has high agreement with the LEDS [127] for long-term contextual threat and demonstrates good concurrent validity with good sensitivity (1.00) and specificity (0.88) for events 3 months prior to data collection [129]. The greater the score on this measure, the more threatening events have occurred. The 12-item LTE-Q was used in this trial.

J. Cognitive Therapy Skills Use

1. Skills of Cognitive Therapy (SoCT)—The 8-item Skills of Cognitive Therapy [SoCT; 130] assesses patients' understanding and use of basic cognitive therapy skills rated from the perspectives of observers (SoCT-O) and patients (SoCT-P). Ratings of patients' skills are made on a 5-point Likert scale ranging from 1 ("never") to 5 ("always or when needed"). The score is the mean of the items; higher scores reflect greater skill in cognitive therapy principles and coping strategies. In this trial, we used a 35-item pool, rated by both patients and their therapists, to evaluate the reliability and validity of both versions of the SoCT and to develop the resulting shorter versions.

K. Treatment Preference

1. Patient Attitudes and Expectations form (PAE)—The Patient Attitudes and Expectations form (PAE) [131] consists of 46 items designed to measure: (a) patients' beliefs

about the cause of their problems, (b) patients' beliefs about helpful treatments, (c) patients' expectations of their therapist, (d) patients' expectations of treatment outcomes, and (e) patients' beliefs about the length of time before treatment is effective. The questions are rated on Likert scales ranging from 1–4 to 1–7. The 46-item PAE form was used in this protocol.

© 2010, Robin B. Jarrett at The University of Texas Southwestern Medical Center at Dallas, all rights reserved.

References

1. First, MB.; Spitzer, RL.; Gibbon, M.; Williams, JB. Structured Clinical Interview for DSM-IV Axis I Disorders-Patient Edition (SCID-I/P, Version 2.0). New York: New York State Psychiatric Institute, Biometrics Research Department; 1996.
2. Riskind JH, Beck AT, Berchick RJ, Brown G, Steer RA. Reliability of the DSM-III diagnoses for major depression and generalized anxiety disorder using the Structured Clinical Interview for DSM-III. *Arch Gen Psychiatry* 1987;44:817–20. [PubMed: 3632255]
3. Keller MB, Shapiro RW, Lavori PW, Wolfe N. Recovery in major depressive disorder: Analysis with the life table and regression models. *Arch Gen Psychiatry* 1982a;39:905–10. [PubMed: 7103679]
4. Keller MB, Shapiro RW, Lavori PW, Wolfe N. Relapse in major depressive disorder: Analysis with the life table. *Arch Gen Psychiatry* 1982b;39(8):911–5. [PubMed: 7103680]
5. Keller MB, Lavori PW, Friedman B, Nielson E, Endicott J, McDonald-Scott P, et al. The Longitudinal Interval Follow-up Evaluation: A comprehensive method for assessing outcome in prospective longitudinal studies. *Arch Gen Psychiatry* 1987;44(6):540–8. [PubMed: 3579500]
6. Warshaw MG, Keller MB, Stout RL. Reliability and validity of the longitudinal interval follow-up evaluation for assessing outcome of anxiety disorders. *J Psychiatr Res* 1994;28:531–45. [PubMed: 7699612]
7. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4. Washington, DC: American Psychiatric Association; 1994.
8. Stewart AL, Hays RD, Ware JE. The MOS short-term general health survey: Reliability and validity in patient populations. *Med Care* 1988;26:724–35. [PubMed: 3393032]
9. Warshaw MG, Dyck I, Allsworth J, Stout RL, Keller MB. Maintaining reliability in a long-term psychiatric study: An ongoing inter-rater reliability monitoring program using the longitudinal interval follow-up evaluation. *J Psychiatr Res* 2001;35:297–305. [PubMed: 11591433]
10. Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatry* 1960;23:56–62. [PubMed: 14399272]
11. Clark LA, Watson D. Tripartite model of anxiety and depression: Psychometric evidence and taxonomic implications. *J Abnorm Psychol* 1991;100:316–36. [PubMed: 1918611]
12. Kobak KA, Brown B, Sharp I, Levy-Mack H, Wells K, Okum F, et al. Sources of unreliability in depression ratings. *J Clin Psychopharmacol* 2009;29(1):82–5. [PubMed: 19142114]
13. Schwab JJ, Bialow MR, Holzer CE. A comparison of two rating scales for depression. *J Clin Psychol* 1967;23:94–6. [PubMed: 6031042]
14. Knesevich JW, Biggs JT, Clayton PJ, Ziegler VE. Validity of Hamilton Rating Scale for Depression. *Br J Psychiatry* 1977;131:49–52. [PubMed: 884416]
15. Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J. An inventory for measuring depression. *Arch Gen Psychiatry* 1961;4:561–71. [PubMed: 13688369]
16. Beck AT, Steer RA, Garbin MG. Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clin Psychol Rev* 1988;8:77–100.
17. Beck, AT.; Brown, GK.; Steer, RA. Beck Depression Inventory-II (BDI-II). San Antonio, TX: The Psychological Corporation; 1996.
18. Rush AJ, Giles DE, Schlessler MA, Fulton CL, Weissenburger JE, Burns CT. The Inventory for Depressive Symptomatology (IDS): Preliminary findings. *Psychiatry Res* 1986;18:65–87. [PubMed: 3737788]
19. Trivedi MH, Rush AJ, Ibrahim HM, Carmody TJ, Biggs MM, Suppes T, et al. The Inventory of Depressive Symptomatology, Clinician rating (IDS-C) and Self-Report (IDS-SR), and the Quick

- Inventory of Depressive Symptomatology (QIDS-C) and Self-Report (QIDS-SR) in public sector patients with mood disorders: A psychometric evaluation. *Psychol Med* 2004;24:73–82. [PubMed: 14971628]
20. Endicott J, Spitzer RL, Fleiss JL, Cohen J. The Global Assessment Scale: A procedure for measuring overall severity of psychiatric disturbance. *Arch Gen Psychiatry* 1976;33:766–71. [PubMed: 938196]
 21. Vatnaland T, Vatnaland J, Friis S, Opjordsmoen S. Are GAF scores reliable in routine clinical use? *Acta Psychiatr Scand* 2006;115(4):326–30. [PubMed: 17355524]
 22. Ramirez C, Ekselius L, Ramklint M. Axis V-Global Assessment of Functioning scale (GAF): Further evaluation of the self-report version. *Eur Psychiatry* 2008;23(8):575–9. [PubMed: 18774271]
 23. Buysse DJ, Reynolds CFI, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: A new instrument for psychiatric practice and research. *Psychiatry Res* 1989;28:193–213. [PubMed: 2748771]
 24. Carpenter JS, Andrykowski MA. Psychometric evaluation of the Pittsburgh Sleep Quality Index. *J Psychosom Res* 1998;45(1):5–13. [PubMed: 9720850]
 25. Kupfer DJ, Frank E, Thase ME. Somatic Symptom Scale. Unpublished measurement.
 26. Peterson C, Semmel A, von Baeyer C, Abramson LY, Matalsky GI, Seligman MP. The Attributional Style Questionnaire. *Cognit Ther Res* 1982;6:287–300.
 27. Peterson C, Villanova P. An Expanded Attributional Style Questionnaire. *J Abnorm Psychol* 1988;97:87–9. [PubMed: 3351118]
 28. Dykema J, Bergbower K, Doctora JD, Peterson C. An Attributional Style Questionnaire for general use. *J Psychoeduc Assess* 1996;14(2):100–8.
 29. Riso LP, Froman SE, Raouf M, Gable P, Maddux RE, Turini-Santorelli N, et al. The long-term stability of early maladaptive schemas. *Cognit Ther Res* 2006;30(4):515–29.
 30. Riso LP, du Toit PL, Blandino JA, Penna S, Dacey S, Duin JS, et al. Cognitive aspects of chronic depression. *J Abnorm Psychol* 2003;112(1):72–80. [PubMed: 12653415]
 31. Bennett KK, Elliott M. Pessimistic explanatory style and cardiac health: What is the relation and the mechanism that links them? *Basic Appl Soc Psych* 2005;27(3):239–48.
 32. Peterson C, Bishop MP, Fletcher CW, Kaplan MR, Yesko ES, Moon CH, et al. Explanatory style as a risk factor for traumatic mishaps. *Cognit Ther Res* 2001;25(6):633–49.
 33. Peterson C, Vaidya RS. Explanatory style, expectations, and depressive symptoms. *Pers Individ Dif* 2001;31:1217–23.
 34. Gladstone TRG, Kaslow NJ. Depression and attributional style in children and adolescents: A meta-analytic review. *J Abnorm Child Psychol* 1995;23:597–606. [PubMed: 8568082]
 35. Reivich, K. The measurement of explanatory style. In: Buchanan, GM.; Seligman, MEP., editors. *Explanatory style*. Hillsdale, NJ: Lawrence Erlbaum; 1995.
 36. Weissman AN. The Dysfunctional Attitudes Scale: A validation study (Doctoral dissertation, University of Pennsylvania, 1979). *Dissertation Abstracts International* 1979;40:1389B–90B.
 37. Otto MW, Teachman BA, Cohen LS, Cohen CN, Vitonis AF, Harlow BL. Dysfunctional attitudes and episodes of major depression: Predictive validity and temporal stability in never-depressed, depressed, and recovered women. *J Abnorm Psychol* 2007;116(3):475–83. [PubMed: 17696703]
 38. Nelson LD, Stern SL, Cicchetti DV. The Dysfunctional Attitudes Scale: How well can it measure depressive thinking? *J Psychopathol Behav Assess* 1992;14(3):217–24.
 39. Beevers CG, Strong DR, Meyer B, Pilkonis PA, Miller IW. Efficiently assessing negative cognition in depression: An item response theory analysis of the Dysfunctional Attitudes Scale. *Psychol Assess* 2007;19(2):199–209. [PubMed: 17563201]
 40. Jarrett RB, Vittengl JR, Clark LA, Dunn TW. Reducing relapse and recurrence in unipolar depression: A comparative meta-analysis of cognitive-behavioral therapy's effects. *J Consult Clin Psychol* 2007;75(3):475–88. [PubMed: 17563164]
 41. Oliver JM, Baumgart EP. The Dysfunctional Attitudes Scale: Psychometric properties and relation to depression in an unselected adult population. *Cognit Ther Res* 1985;9(2):161–7.
 42. Beevers CG, Keitner GI, Ryan CE, Miller IW. Cognitive predictors of symptom return following depression treatment. *J Abnorm Psychol* 2003;112(3):488–96. [PubMed: 12943027]

43. Segal ZV, Kennedy S, Gemar M, Hood K, Pedersen R, Buis T. Cognitive reactivity to sad mood provocation and the prediction of depressive relapse. *Arch Gen Psychiatry* 2006;63(7):749–55. [PubMed: 16818864]
44. Dobson KS, Breiter HJ. Cognitive assessment of depression: Reliability and validity of three measures. *J Abnorm Psychol* 1983;92:107–9. [PubMed: 6833629]
45. Haefffel GJ, Abramson LY, Voelz ZR, Metalsky GI, Halberstadt L, Dykman BM, et al. Negative cognitive styles, dysfunctional attitudes, and the remitted depression paradigm: A search for the elusive cognitive vulnerability to depression factor among remitted depressives. *Emotion* 2005;5(3):343–8. [PubMed: 16187869]
46. Hankin BL, Fraley RC, Abela JR. Daily depression and cognitions about stress: Evidence for a traitlike depressogenic cognitive style and the prediction of depressive symptoms in a prospective daily diary study. *J Pers Soc Psychol* 2005;88(4):673–85. [PubMed: 15796667]
47. Ilardi SS, Craighead WE. The relationship between personality pathology and dysfunctional cognitions in previously depressed adults. *J Abnorm Psychol* 1999;108(1):51–7. [PubMed: 10066992]
48. Hamilton EW, Abramson LY. Cognitive patterns and major depressive disorders: A longitudinal study in a hospital setting. *J Abnorm Psychol* 1983;92:173–84. [PubMed: 6863732]
49. McNair, DM.; Lorr, M.; Droppleman, LF. EdITS manual for the Profile of Mood States. San Diego, CA: EdITS; 1992.
50. Floyd M, Scogin F, Chaplin WF. The Dysfunctional Attitudes Scale: Factor structure, reliability, and validity with older adults. *Aging Ment Health* 2004;8(2):153–60. [PubMed: 14982720]
51. Esther de Graaf L, Roelofs J, Huibers MJH. Measuring dysfunctional attitudes in the general population: The Dysfunctional Attitude Scale (Form A) Revised. *Cognitive Therapy and Research*. in press.
52. Chioqueta AP, Stiles TC. Psychometric properties of the Norwegian version of the Dysfunctional Attitude Scale (Form A). *Cogn Behav Ther* 2004;33(2):83–6. [PubMed: 15279314]
53. Tanaka N, Uji M, Hiramura H, Chen Z, Shikai N, Kitamura T. Cognitive patterns and depression: Study of a Japanese university student population. *Psychiatry Clin Neurosci* 2006;60:358–64. [PubMed: 16732754]
54. Wong DFK, SCK, Lau Y. The reliability and validity of the Chinese version of the Dysfunctional Attitudes Scale Form A (DAS-A) in a community sample. *Int J Psychiatry Med* 2008;38(2):141–52. [PubMed: 18724566]
55. Beck AT, Weissman A, Lester D, Trexler L. The measurement of pessimism: The Hopelessness Scale. *J Consult Clin Psychol* 1974;42:861–5. [PubMed: 4436473]
56. Beck, AT.; Steer, RA. Manual for the Beck Hopelessness Scale. San Antonio, TX: The Psychological Corporation; 1988.
57. Minkoff K, Bergman E, Beck AT, Beck R. Hopelessness, depression, and attempted suicide. *Am J Psychiatry* 1973;130:455–9. [PubMed: 4691303]
58. Beck AT, Kovacs M, Weissman A. Hopelessness and suicidal behavior: An overview. *JAMA* 1975;234:1146–9. [PubMed: 1242427]
59. Kazdin AE, French NH, Unis AS, Esveldt-Dawson K, Sherick RB. Hopelessness, depression, and suicidal intent among psychiatrically disturbed inpatient children. *J Consult Clin Psychol* 1983;51:504–10. [PubMed: 6619356]
60. Beck AT, Steer RA, Kovacs M, Garrison BS. Hopelessness and eventual suicide: A 10-year prospective study of patients hospitalized with suicidal ideation. *Am J Psychiatry* 1985;142:559–63. [PubMed: 3985195]
61. McMillan D, Gilbody S, Beresford E, Neilly L. Can we predict suicide and non-fatal self-harm with the Beck Hopelessness Scale? A meta-analysis *Psychol Med* 2007;37:769–78.
62. Beck AT, Steer R, Trexler LD. Alcohol abuse and eventual suicide: A 5- to 10-year prospective study of alcohol-abusing suicide attempters. *J Stud Alcohol* 1989;50(3):202–9. [PubMed: 2724967]
63. Haatainen K, Tanskanen A, Kylma J, Honkalampi K, Koivumaa-Honkanen H, Hintikka J, et al. Factors associated with hopelessness: A population study. *Int J Soc Psychiatry* 2004;50(2):142–52. [PubMed: 15293431]

64. Steed L. Further validity and reliability evidence for Beck Hopelessness Scale scores in a nonclinical sample. *Educ Psychol Meas* 2001;61:303–16.
65. Rosenbaum M. A schedule for assessing self-control behaviors: Preliminary findings. *Behav Ther* 1980;11:109–21.
66. Akgun S. The effects of situation and learned resourcefulness on coping responses. *Soc Behav Pers* 2004;32(5):441–8.
67. Bernard L. Individual differences in vigor and deliberation: Development of two new measures from an evolutionary psychology theory of human motivation. *Psychol Rep* 2008;103:245–70.
68. Mezo PG, Heiby EM. A comparison of four measures of self-control skills. *Assessment* 2004;11(3):238–50. [PubMed: 15358880]
69. Slessareva E, Muraven M. Sensitivity to punishment and self-control: The mediating role of emotion. *Pers Individ Dif* 2004;36:307–19.
70. Jarrett RB, Giles D, Gullion C, Rush A. Does learned resourcefulness predict response to cognitive therapy in depressed outpatients? *J Affect Disord* 1991;23(4):223–9. [PubMed: 1791268]
71. Burns DD, Rude S, Simons AD, Bates MA, Thase ME. Does learned resourcefulness predict the response to cognitive behavioral therapy for depression? *Cognit Ther Res* 1994;18:277–91.
72. Bekhet A, Zauszneiwski JA, Wykle ML. Milieu change and relocation adjustment in elders. *J Nurs Res* 2008;30(1):113–29.
73. Ngai F, Chan SW, Holroyd E. Translation and validation of a Chinese version of the Self-Control Schedule in Chinese childbearing women. *J Clin Nurs* 2008;17:323–32. [PubMed: 17419785]
74. Boonpongmanee C, Zauszniewski JA, Boonpongmanee S. Psychometric properties of the Self-Control Schedule: Thai version. *Res Nurs Health* 2002;25(6):471–8. [PubMed: 12424784]
75. Gemar MC, Segal ZV, Sagrati S, Kennedy SJ. Mood-induced changes on the Implicit Association Test in recovered depressed patients. *J Abnorm Psychol* 2001;10(2):282–9. [PubMed: 11358022]
76. Bowen R, Clark M, Baetz M. Mood swings in patients with anxiety disorders compared with normal controls. *J Affect Disord* 2004;78:185–92. [PubMed: 15013242]
77. Lethbridge R, Allen M. Mood-induced cognitive and emotional reactivity, life stress, and the prediction of depressive relapse. *Behav Res Ther* 2008;46:1142–50. [PubMed: 18707676]
78. Marzillier SL, Davey GCL. Anxiety and disgust: Evidence for a unidirectional relationship. *Emotion* 2005;19(5):729–50.
79. Scherrer MC, Dobson KS. Mood-induced cognitive and emotional reactivity, life stress, and the prediction of depressive relapse. *J Clin Psychol* 2009;65(1):20–35. [PubMed: 19051242]
80. Grossberg JM, Grant BF. Clinical psychophysics: Applications of ratio scaling and signal detection methods to research on pain, fear, drugs, and medical decision making. *Psychol Bull* 1978;85(5):1154–76. [PubMed: 704721]
81. Williams ACDC, Davies HTO, Chadury Y. Simple pain rating scales hide complex idiosyncratic meanings. *Pain* 2000;85(3):457–63. [PubMed: 10781919]
82. Hofmans J, Theuns P. On the linearity of predefined and self-anchoring Visual Analogue Scales. *Br J Math Stat Psychol* 2008;61(2):401–13. [PubMed: 17535485]
83. Hofmans J, Theuns P, van Acker F. Combining quality and quantity: A psychometric evaluation of the self-anchoring scale. *Qual Quant* 2009;43(5):703–16.
84. Spanier GB. Measuring dyadic adjustment: New scales for assessing the quality of marriage and similar dyads. *J Marriage Fam* 1976;38:15–28.
85. Graham JM, Liu YJ, Jeziorski JL. The Dyadic Adjustment Scale: A reliability generalization meta-analysis. *J Marriage Fam* 2006;68:701–17.
86. Crane DR, Busby DM, Larson JH. A factor analysis of the Dyadic Adjustment Scale with distressed and nondistressed couples. *Am J Fam Ther* 1991;19:60–6.
87. Spanier GB. Assessing the strengths of the Dyadic Adjustment Scale. *J Fam Psychol* 1988;2:92–4.
88. Locke HJ, Wallace KM. Short marital adjustment and prediction tests: Their reliability and validity. *Marriage Fam Living* 1959;21:251–5.
89. Weissman MM, Bothwell S. Assessment of social adjustment by patient self-report. *Arch Gen Psychiatry* 1976;33(9):1111–5. [PubMed: 962494]

90. Weissman MM, Prusoff BA, Thompson D, Harding PS, Myers JK. Social adjustment by self-report in a community sample and in psychiatric outpatients. *J Nerv Ment Dis* 1978;166:317–26. [PubMed: 650195]
91. Raskin A, Schulterbrandt JG, Reatig N, McKeon JJ. Replication of factors of psychopathology in interview, ward behavior, and self-report ratings of hospitalized depressives. *J Nerv Ment Dis* 1969;148:87–98. [PubMed: 5768895]
92. Markush, RE.; Favero, RV. Epidemiologic assessment of stressful life events, depressed mood and psychophysiological symptoms: A preliminary report. In: Dohrenwend, BS.; Dohrenwend, BP., editors. *Stressful life events: Their nature and effects*. New York: John Wiley & Sons; 1974.
93. Derogatis LR, Lipman RS, Covi L. The SCL-90: An outpatient psychiatric rating scale. *Psychopharmacol Bull* 1973;9:13–29. [PubMed: 4682398]
94. Suzuki Y, Sakurai A, Yasuda T, Harai H, Kitamura T, Takahashi K, et al. Reliability, validity, and standardization of the Japanese version of the Social Adjustment Scale-Self Report. *Psychiatry Clin Neurosci* 2003;57(4):441–6. [PubMed: 12839527]
95. Zweig RA, Turkel E. The Social Adjustment Scale-Self-Report: Psychometric properties for older adults. *Psychol Rep* 2007;101(3):920–6. [PubMed: 18232450]
96. King S, Dixon MJ. Expressed emotion, family dynamics and symptom severity in a predictive model of social adjustment for schizophrenic young adults. *Schizophr Res* 1995;14:121–32. [PubMed: 7710992]
97. Schooler, N.; Hogarty, G.; Weissman, MM. Social Adjustment Scale II (SAS II). In: Hargreaves, WA.; Attkisson, CC.; Sorenson, JE., editors. *Resource materials for community mental health program evaluators*. Rockville, MD: U.S. Department of Health, Education, and Welfare; 1979. p. 290-303.
98. Weiduwilt KM, Jerrell JM. The reliability and validity of the SAS-SMI. *J Psychiatr Res* 1999;33:105–12. [PubMed: 10221742]
99. Horowitz LM, Rosenberg SE, Baer BA, Ureno G, Villasenor VS. Inventory of Interpersonal Problems: Psychometric properties and clinical applications. *J Consult Clin Psychol* 1988;56(6):885–92. [PubMed: 3204198]
100. Woodward LE, Murrell SA, Bettler RF. Stability, reliability, and norms for the Inventory of Interpersonal Problems. *Psychother Res* 2005;15(3):272–86.
101. Barkham M, Hardy GE, Startup M. The structure, validity, and clinical relevance of the Inventory of Interpersonal Problems. *Br J Med Psychol* 1994;67:171–85. [PubMed: 7918212]
102. Derogatis, LR. *The SCL-90-R: Administration, scoring, and procedures - Manual II*. Towson, MD: Clinical Psychometric Research; 1983.
103. Gurtman MB. Interpersonal problems and the psychotherapy context: The construct validity of the Inventory of Interpersonal Problems. *Psychol Assess* 1996;8:241–55.
104. Hooley J, Teasdale J. Predictors of relapse in unipolar depressives: Expressed emotion, marital distress, and perceived criticism. *J Abnorm Psychol* 1989;98(3):229–35. [PubMed: 2768657]
105. Leff, J.; Vaughn, C. *Expressed emotion in families: Its significance for mental illness*. New York: Guilford Press; 1985.
106. Wing, JK.; Cooper, JE.; Sartorius, N. *The measurement and classification of psychiatric symptoms*. Cambridge: Cambridge University Press; 1974.
107. Riso LP, Klein DN, Anderson RL, Ouimette PC, Lizardi H. Convergent and discriminant validity of perceived criticism from spouses and family members. *Behav Ther* 1996;27(2):129–37.
108. Derogatis, LR. *SCL-90-R: Administration, scoring, and procedures manual*. 3. Towson, MD: Clinical Psychometric Research; 1994.
109. Chambless DL, Bryan AD, Aiken LS, Steketee G, Hooley JM. Predicting expressed emotion: A study with families of obsessive-compulsive and agoraphobic outpatients. *J Fam Psychol* 2001;15(2):225–40. [PubMed: 11458630]
110. Zimmerman M, Coryell W, Corenthal C, Wilson S. A self-report scale to diagnose Major Depressive Disorder. *Arch Gen Psychiatry* 1986;43(11):1076–81. [PubMed: 3767597]
111. Hooley JM, Parker HA. Measuring expressed emotion: An evaluation of the shortcuts. *J Fam Psychol* 2006;20(3):386–96. [PubMed: 16937995]

112. Horvath AO, Greenberg LS. Development and validation of the Working Alliance Inventory. *J Couns Psychol* 1989;36:223–33.
113. Bordin, ES. Of human bonds that bind or free. Presidential address delivered at the meeting of the Society for Psychotherapy Research; Pacific Grove, CA. 1980, June;
114. Howard I, Turner R, Olkin R, Mohr DC. Therapeutic alliance mediates the relationship between interpersonal problems and depression outcome in a cohort of multiple sclerosis patients. *J Clin Psychol* 2006;62(9):1179–204.
115. Martin DJ, Garske JP, Davis MK. Relation of the therapeutic alliance with outcome and other variables: A meta-analytic review. *J Consult Clin Psychol* 2000;68(3):438–50. [PubMed: 10883561]
116. Clark, LA. Manual for the Schedule for Nonadaptive and Adaptive Personality (SNAP). Minneapolis, MN: University of Minnesota Press; 1993.
117. Clark, LA.; Simms, LJ.; Wu, KD.; Casillas, A. The Schedule for Nonadaptive and Adaptive Personality, 2nd Edition (SNAP-2). Minneapolis, MN: University of Minnesota Press; in press
118. Eysenck, HJ.; Eysenck, SG. Eysenck Personality Questionnaire Manual. San Diego, CA: Educational and Industrial Testing Service; 1975.
119. Zuckerman M, Kuhlman DM, Camac C. What lies beyond E and N? Factor analysis of scales believed to measure basic dimensions of personality. *J Pers Soc Psychol* 1988;54:96–107.
120. Watson D, Clark LA, Harkness AR. Structures of personality and their relevance to psychopathology. *J Abnorm Psychol* 1994;103:18–31. [PubMed: 8040477]
121. Pfohl, B.; Blum, N.; Zimmerman, M. Structured interview for DSM-IV Personality. Washington, D.C: American Psychiatric Press; 1997.
122. Endicott J, Nee J, Harrison W, Blumenthal R. Quality of Life Enjoyment and Satisfaction Questionnaire: A new measure. *Psychopharmacol Bull* 1993;29(2):321–6. [PubMed: 8290681]
123. Ritsner M, Kurs R, Gibel A, Ratner Y, Endicott J. Validity of an abbreviated Quality of Life Enjoyment and Satisfaction Questionnaire (Q-LES-Q-18) for schizophrenia, schizoaffective, and mood disorder patients. *Qual Life Res* 2005;14(7):1693–703. [PubMed: 16119181]
124. Endicott J, Nee J, Yang R, Wolberg C. Pediatric Quality of Life Enjoyment and Satisfaction Questionnaire (PQ-LES-Q): Reliability and validity. *J Am Acad Child Adolesc Psychiatry* 2006;45(4):401–7. [PubMed: 16601644]
125. Rossi A, Rucci P, Mauri M, Maina G, Pieraccini F, Pallanti S, et al. Validity and reliability of the Italian version of the Quality of Life Enjoyment and Satisfaction Questionnaire. *Qual Life Res* 2005;14(10):2323–8. [PubMed: 16328911]
126. Brugha TS, Bebbington P, Tennant C, Hurry J. The List of Threatening Experiences: A subset of 12 life event categories with considerable long-term contextual threat. *Psychol Med* 1985;15:189–94. [PubMed: 3991833]
127. Brown, GW.; Harris, T. Social origins of depression. London: Tavistock Publications; 1978.
128. Tennant C, Andrews G. A scale to measure life events. *Aust N Z J Psychiatry* 1976;10:27–32. [PubMed: 1065311]
129. Brugha TS, Cragg D. The List of Threatening Experiences: The reliability and validity of a brief life events questionnaire. *Acta Psychiatr Scand* 1990;82:77–81. [PubMed: 2399824]
130. Jarrett, RB.; Vittengl, JR.; Thase, ME.; Clark, LA. Skills of Cognitive Therapy (SoCT): A new measure of patients' comprehension and use. In preparation
131. Sotsky SM, Glass DR, Shea TS, Pilkonis PA, Collins JF, Elkin I, et al. Patient predictors of response to psychotherapy and pharmacology: Findings in the NIMH Treatment of Depression Collaborative Research Program. *Am J Psychiatry* 1991;148:997–1008. [PubMed: 1853989]

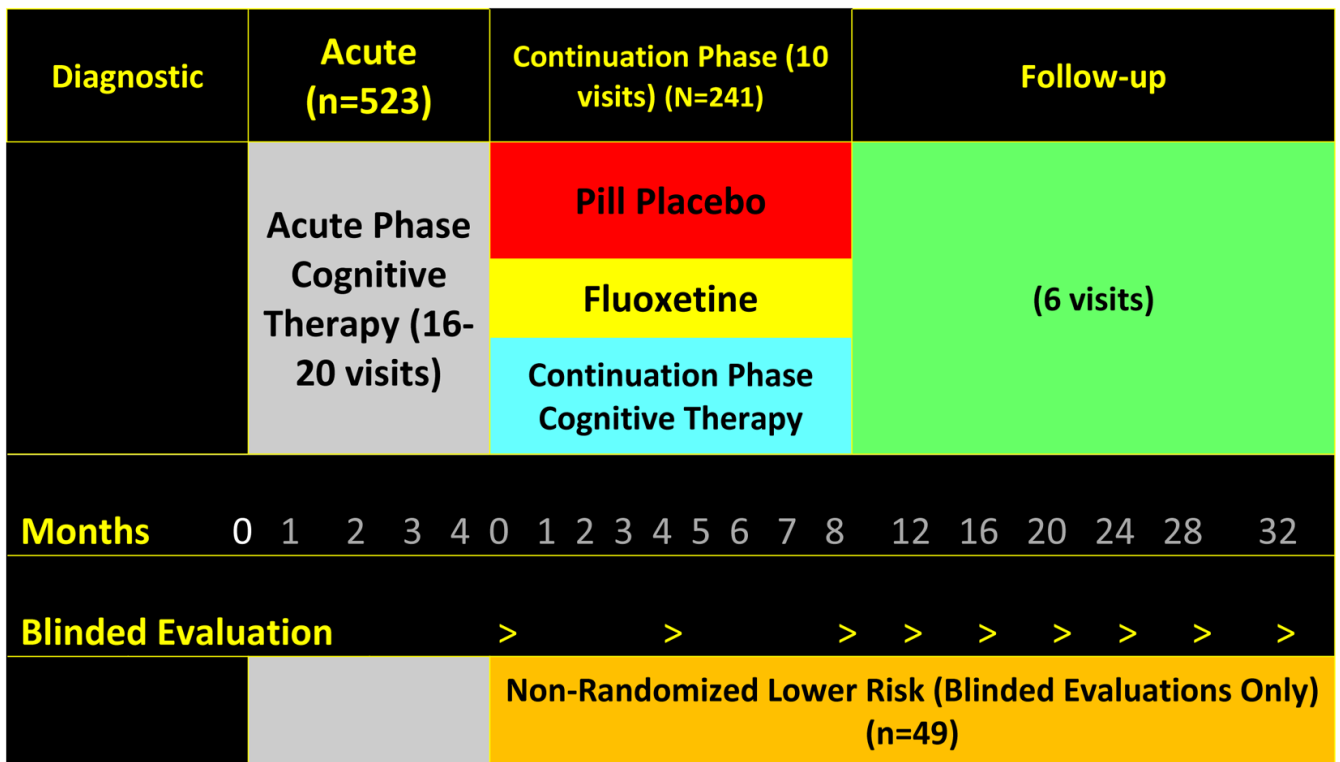


Figure 1. Study Design

This is a figure of the research design detailing the study phases, time periods, and number of participants.

© 2010, Robin B. Jarrett at The University of Texas Southwestern Medical Center at Dallas, all rights reserved.

Table 1

Phases and Timing of Protocol Events

| Phase | Months in study | Time | Week | Event |
|--|-----------------|-------|-------|--|
| Diagnostic Evaluation | 0 | T1 | -1 | Triage Evaluation; Psychiatric Evaluation; medical screen and drug wash-out; Informed Consent; 1st Psychoeducational Session |
| Acute Phase-Cognitive Therapy (approx. 12-14 weeks) | | | | Sessions |
| | | | 1 | 1-2 |
| | | | 2 | 3-4 |
| | | | 3 | 5-6 |
| | | | 4 | 7-8 |
| | 1 | | | HRSD-17 reduction from baseline |
| | | | 5 | 9 |
| | | | 6 | 10 |
| | | | 7 | 11 |
| | | | 8 | 12 |
| | 2 | | | 13 |
| | | 9 | 14 | |
| | | 10 | 15 | |
| | | 11 | 16 | |
| 3+ | | | 17 | |
| | | 12-14 | 18 | |
| | | | 19 | |
| | | | 20 | |
| Blinded Evaluation ** Randomization | 4 | T2 | 13-16 | 7 |
| Continuation Phase (8 months) | | | | HRSD-17 reduction from baseline |
| | | | | ≥40%(early responders) |
| | | | | <40%(late responders) |
| | | | | HRSDs collected for Risk Assessment |
| | | | | HRSDs collected for Risk Assessment |
| | | | | HRSDs collected for Risk Assessment |
| | | | | HRSDs collected for Risk Assessment |
| | | | | HRSDs collected for Risk Assessment |
| | | | | HRSDs collected for Risk Assessment |
| | | | | HRSDs collected for Risk Assessment |
| Blinded Evaluation ** | | | | Higher Risk |
| | | | | Lower Risk |
| | | | | Higher Risk |
| | | | | Lower Risk |
| | | | | Higher Risk |
| | | | | Lower Risk |

Table 2

Data Collection (T1-T10)

| Measurement Occasion | T1 | | | | | T2 | | | | | T3-T10 | |
|---|-----------------------|-----------|-------------|---------------------|-------------|--------------------|--------------|---------------------|---|------|--------|-------|
| | Diagnostic Evaluation | Follow-Up | Acute Phase | Educational Session | BEV Random. | Continuation Phase | Clinic Value | Blinded Evaluations | | | | |
| <u>Months from Randomization</u> | -4 | -4 | -3 | -1 | 0 | 1-8 | | | | 4-32 | | |
| Demographics | | | | | | | | | | | | |
| Patient Background (BKGD) | X | | | | | | | | | | | |
| Psychiatric Diagnosis | | | | | | | | | | | | |
| Structured Clinical Interview for DSM-IV (SCID) | X | | | | X | | | | X | | | X |
| Current MDE (Section A) alone | X | | X(1) | | X | X(2) | | | X | | | X |
| Longitudinal Interval Follow-up Evaluation | | | | | | | | | | | | |
| Psychiatric Status Rating (LIFE PSR) | | | | | X | | | | | | | X |
| MDE Specifiers and Past MDE (Past MDE) | X | | | | | | | | | | | |
| Initial Evaluation Summary (IES) | X | | | | | | | | | | | |
| Symptom Severity | | | | | | | | | | | | |
| Hamilton Rating Scale for Depression (HRSD) | X | X(6) | X(7) | | X | X(2) | | | X | | | X |
| Beck Depression Inventory (BDI) | X | | X(7) | | X | X(2) | | | X | | | X |
| Global Assessment of Functioning (GAF) | X | X(6) | X(7) | | X | X(2) | | | X | | | X |
| Inventory of Depressive Symptomatology | | | | | | | | | | | | |
| Self-Report (IDS-SR) | X | | X(7) | | X | X(2) | | | X | | | X |
| Pittsburgh Sleep Quality Index (PSQI) | | X | X(8,1) | | | X(9) | | | X | | | X(17) |
| Somatic Symptoms Scale (SSS) | | | | | | X(2) | | | X | | | |
| Personality | | | | | | | | | | | | |
| Schedule for Normal and Abnormal Personality (SNAP) | | | X(3) | | | X(4) | | | | | | X(5) |
| Cognitive | | | | | | | | | | | | |
| Attributional Style Questionnaire (ASQ) | X | | | X | X | X(10) | | | | | | X |

| Measurement Occasion | T1 | T2 | T3-T10 |
|---|----|-------|--------|
| Beck Hopelessness Scale (BHS) | | X | X |
| Self Control Schedule (SCS) | X | X | X |
| Dysfunctional Attitudes Scale (DAS) | | | |
| Form A | X | X | X(10) |
| Form B | | X(11) | X(12) |
| Mood Induction | | X(11) | X(12) |
| Interpersonal Functioning | | | |
| Social Adjustment Scale (SAS-SR) | X | X | X(10) |
| Dyadic Adjustment Scale (DYS) | X | X | X(10) |
| Inventory of Interpersonal Problems (IIP) | X | X | X(10) |
| Working Alliance Inventory | | | |
| Patient Version (WAIC) | | X | X(16) |
| Therapists Version (WAIT) | | | X(14) |
| Perceived Criticism Scale (PCS) | X | X | X |
| Longitudinal Interval Follow-up Evaluation | | | |
| Psychosocial Functioning (LIFE PSF) | | X | X |
| Quality of Life | | | |
| Quality of Life Enjoyment and Satisfaction Questionnaire (QOLESQ) | X | X | X |
| Life Events | | | |
| List of Threatening Events (LTE) | X | X | X |
| Cognitive Therapy Skills Use | | | |
| Skills of Cognitive Therapy | | | |
| Patient Version (SoCT-P) | | X | X |
| Observer Version (SoCT-O) | | | X(14) |
| Treatment Preference | | | |
| Patient Attitudes & Expectations Scale (PAES) | X | | X(15) |
| Non-Protocol Treatment | | | |

| Measurement Occasion | T1 | T2 | T3-T10 |
|--|----|----|--------|
| Longitudinal Interval Follow-up Evaluation Treatment History (LIJE TXHX) | | X | X |
| 1. Acute Phase CT weeks 4, 8, 12 | | | |
| 2. Continuation Phase sessions 1–10 (semi-monthly for months 1–2; monthly for months 3–8) | | | |
| 3. Acute Phase CT weeks 1 and 12 | | | |
| 4. Continuation phase session 10 | | | |
| 5. Month 32 or Exit | | | |
| 6. Repeated at least 5 business days after initial evaluation and within 5 business days of the first Acute Phase CT session | | | |
| 7. Once weekly during Acute Phase CT | | | |
| 8. Acute Phase CT Session 1 | | | |
| | | | T=Time |

Appendix II

Contributors to Research

| UT Southwestern Medical Center at Dallas | University of Pittsburgh |
|--|--------------------------|
| <i>Research Staff</i> | <i>Research Staff</i> |
| Amy Bailey, M.S. | Susan Berman, M.Ed. |
| Monica Basco, Ph.D. | Judith S. Callan, Ph.D. |
| Patricia Borman, Ph.D. | Trisha Defrances, B.S. |
| Anthony Cordaro, M.D. | Tim Harden, B.A. |
| Nancy Cravens, L.V.N. | Jim Moorehead, B.S. |
| Heidi Damin, M.Ed. | Crystal R. Spotts, M.Ed. |
| Kimberly Doyle, Ph.D. | Luann Shutt, M.S.N. |
| Lauren Dunlap, B.A. | Greg Siegle, Ph.D. |
| Todd Dunn, Ph.D. | Lisa Stupar |
| Barbara Foster, Ph.D. | |
| Janine Rodenhiser-Hill, Ph.D. | |
| Shelley Guess, B.A. | |
| Agnes Hamilton, M.L.A. | |
| Margeaux Hollenbeck, B.A. | |
| Nisa Holzapfel, M.S. | |
| Nahid Hooshyar, Ph.D. | |
| Bruce Jones, M.A. | |
| Jessica Jones, M.A. | |
| Julie Kangas, B.A. | |
| Cindy L. Kidner, Ph.D. | |
| Christine Killebrew, B.A. | |
| Andrew Lumpkin, B.S. | |
| Kevin Majeres, M.D. | |
| Abu Minhajuddin, Ph.D. | |
| Paula Olive, B.S.W. | |
| Elizabeth Powers-Wade, B.A. | |
| Joanne Sanders, M.S. | |
| Sharon Sowell, B.S. | |
| Daniel Taylor, Ph.D. | |
| Margaret K. Wagman, M.S. | |
| Heather M. Walters, M.S. | |
| Michelle White, B.S. | |
| Marjorie Woodruff, Ph.D. | |
| Numerous work-study students | |
| <i>Therapists</i> | <i>Therapists</i> |
| Gregory Eaves, Ph.D. | Claudia DiNardo, M.Ed. |

| UT Southwestern Medical Center at Dallas | University of Pittsburgh |
|---|---------------------------------|
| William L. Edwards, Ph.D. | Sander Kornblith, Ph.D. |
| Bethany R. Hampton, Ph.D. | Maureen Maher Bridge, M.S.W. |
| Robin B. Jarrett, Ph.D. | Elizabeth Pacoe, Ph.D. |
| Dolores Kraft, Ph.D. | Tiffany Painter, M.S.W. |
| Alexandra Ramos, Ph.D. | Mary Ponzer, M.D. |
| Paul Silver, Ph.D. | Elizabeth Strayhorn, M.S.W. |
| | Karen Woodall, Ph.D. |
| <i>Blinded Evaluators</i> | <i>Blinded Evaluators</i> |
| Jeanette Doyle, M.A. | Michele Feingold, M.S.W. |
| Berit Johnson, Ph.D. | Rebecca Hartjen, M.S.W. |
| Cathy Judd, P.S.-C, M.S. | Lydia Lechlitter, Ph.D. |
| Steve Krebaum, Ph.D. | Michael Lightfoot, M.S. |
| Maria Marwill-Magee, Ph.D. | |
| Mark Williams, Ph.D. | |
| <i>Psychopharmacologists</i> | <i>Psychopharmacologists</i> |
| Wayne Denton, M.D., Ph.D. | Robert H. Howland, M.D. |
| Howard K. Gershenfeld, M.D., Ph.D | Edward S. Friedman, M.D. |
| | Duane Spiker, M.D. |
| | Timothy Denko, M.D. |
| | Michael E. Thase, M.D. |
| <i>Pharmacists</i> | <i>Pharmacists</i> |
| Paul Nguyen, R.Ph. | Christine Nguyen, R.Ph. |
| Michael Phan, Pharm.D. | |
| <i>Data Safety and Monitoring Board Members</i> | |
| Edward S. Friedman, M.D. | |
| Rohan Ganguli, M.D. | |
| Joel Greenhouse, Ph.D. | |
| Roger Haskett, M.D. | |
| Beverley Huet, Ph.D. | |
| Sheryl Kelsey, Ph.D. | |
| Jane Pearson, Ph.D. | |
| Linda Street, Ph.D. | |