



Published in final edited form as:

J Consult Clin Psychol. 2013 October ; 81(5): 783–792. doi:10.1037/a0033250.

Detecting Critical Decision Points in Psychotherapy and Psychotherapy + Medication for Chronic Depression

Dana Steidtmann, Ph.D.^{1,2}, Rachel Manber, Ph.D.², Christine Blasey, Ph.D.², John C. Markowitz, M.D.³, Daniel N. Klein, Ph.D.⁴, Barbara O. Rothbaum, Ph.D.⁵, Michael E. Thase, M.D.⁶, James H. Kocsis, M.D.⁷, and Bruce A. Arnow, Ph.D.²

²Department of Psychiatry & Behavioral Sciences, Stanford University

³Columbia University/New York State Psychiatric Institute

⁴Department of Psychology, Stony Brook University

⁵Emory University School of Medicine

⁶University of Pennsylvania School of Medicine/Philadelphia Veterans Affairs Medical Center

⁷Department of Psychiatry, Weill Cornell Medical College

Abstract

Objective—To quantify clinical decision points for identifying depression treatment non-remitters prior to end-of-treatment.

Method—Data come from the psychotherapy arms of a randomized clinical trial for chronic depression. Participants ($n=352$; 65.6% female; 92.3% White; mean age = 44.3 years) received 12 weeks of Cognitive Behavioral Analysis System of Psychotherapy (CBASP) or CBASP plus an antidepressant medication. In half of the sample, receiver operating curve (ROC) analyses were used to identify efficient percent symptom reduction cut points on the Inventory of Depressive Symptoms-Self Report (IDS-SR) for predicting end-of-treatment nonremission based on the Hamilton Rating Scale for Depression (HRSD). Sensitivity, specificity, predictive values and Cohen's kappa for identified cut points were calculated using the remaining half of the sample.

Results—Percent IDS-SR symptom reduction at weeks 6 and 8 predicted end of treatment HRSD remission status in both the combined treatment (week 6 cut point = 50.0%, Cohen's kappa = .42; week 8 cut point = 54.3%, Cohen's kappa = .45), and psychotherapy only (week 6 cut point = 60.7%, Cohen's kappa = .41; week 8 cut point = 48.7%, Cohen's kappa = .49). Week 8 was more reliable for identifying nonremitters in psychotherapy only treatment.

¹Corresponding author. Department of Psychiatry and Behavioral Sciences, Stanford University, 401 Quarry Rd., Stanford, CA 94305-5722, dsteidt@stanford.edu.

Dr. Thase has no conflicts of interest pertaining to this paper, although he does report the following relationships with companies that develop treatment for depression or provide education pertaining to those treatments:

Dr. Thase has provided scientific consultation to Alkermes, Astra-Zeneca, Bristol-Myers Squibb Company, Dey Pharma, L.P., Eli Lilly & Company, Forest Pharmaceuticals, Inc., Gerson Lehman Group, GlaxoSmithKline, Guidepoint Global, H. Lundbeck A/S, MedAvante, Inc., Merck and Co. Inc., Neuronetics, Inc., Novartis, Otsuka, Ortho-McNeil Pharmaceuticals, PamLab, L.L.C., Pfizer (formerly Wyeth-Ayerst Laboratories), Schering-Plough (formerly Organon, Inc.), Shire US Inc., Sunovion Pharmaceuticals, Inc., Takeda (Lundbeck), and Transcept Pharmaceuticals. Dr. Thase receives grant funding from the Agency for Healthcare Research and Quality, Eli Lilly & Company, GlaxoSmithKline (ended 7/10), National Institute of Mental Health, Otsuka Pharmaceuticals, and Sepracor, Inc. He has equity holdings in MedAvante, Inc. and receives royalty income from American Psychiatric Foundation, Inc., Guilford Publications, Herald House, Oxford University Press, and W.W. Norton & Company. His wife is employed as the Group Scientific Director for (Embryon – formerly Advogent; which does business with BMS and Pfizer/Wyeth).

Conclusions—Those with chronic depression who will not remit in structured, time-limited psychotherapy for depression, either alone or in combination with antidepressant medication, are identifiable prior to end-of-treatment. Findings provide an operationalized strategy for designing adaptive psychotherapy interventions.

Keywords

chronic depression; psychotherapy for depression; psychotherapy nonremission; adaptive designs; receiver operating curve

Introduction

Despite evidence documenting the effectiveness of psychotherapy, antidepressant medications and combined treatment approaches for depression (Cuijpers, Andersson, Donker, & van Straten, 2011); (Spielmans, Berman, & Usitalo, 2011); (Cuijpers, van Straten, Warmerdam, & Andersson, 2009), these “first line” treatments have modest outcomes. For example, in large randomized clinical trials of evidence-based psychotherapies for depression, only 25–45% of people showed remission of symptoms (Arnow & Hill, 2008). Remission rates have been similarly low in large-scale antidepressant treatment trials using a single antidepressant medication (Fava et al., 2003). Among those who do respond to treatments for major depression, relapse and recurrence rates are high (see Richards, 2011 for a review).

Murphy et al. (2007) suggest that the combination of low remission rates with high relapse and recurrence rates makes the use of sequential treatment decision-making strategies a particularly useful approach for depression treatment (Murphy, Oslin, Rush, & Zhu, 2007). These strategies, frequently referred to as stepped care, expert systems or treatment algorithms, can be evaluated in sequential multiple assignment randomized trials (SMART) (Almirall, Compton, Gunlicks-Stoessel, Duan, & Murphy, 2012; Murphy, 2005). Such strategies may optimize treatment outcomes by providing a framework for clinical decision points based on the observed effects of prior treatment (Lavori & Dawson, 2000). Identifying people who are unlikely to remit could prompt modification of their treatment, and hence would likely improve treatment outcomes. Identifying specific critical treatment decision points is a necessary first step for developing adaptive strategies in order to determine when a treatment change is indicated (Murphy et al., 2007).

Indeed, the development of pharmacotherapy algorithms for major depression grew in part from data on expected time to response for antidepressant medications (Quitkin et al., 1996). Such algorithms specify decision points and accompanying changes in “treatment tactics” (Crismon et al., 1999 (Lavori, Dawson, & Rush, 2000). For example, in the Texas Medication Algorithm Project (TMAP) (Crismon et al., 1999) and the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) trial (Rush et al., 2004; Rush, Trivedi, et al., 2006), researchers developed antidepressant algorithms basing the prescription of medication(s) for a given course of treatment on antidepressant treatment history, and basing dosing, medication switches and medication augmentations on treatment response and tolerability of side effects. For example, the TMAP protocol specifies that by week 4 of treatment, the clinician should increase the medication dose for a patient who has received a therapeutic dose of an agent, has no dose-limiting side effects, and has responded with less than 25% improvement (Crismon et al., 1999). TMAP and other studies using algorithm driven medication management showed improved symptom outcomes compared to medication treatment-as-usual (Trivedi et al., 2004) and single antidepressant trials (Rush, Trivedi, et al., 2006) although some have questioned the magnitude of the accrued benefit (Pigott, Leventhal, Alter, & Boren, 2010). Recognizing the need for complex treatment

approaches, the National Institute of Mental Health (NIMH) has called for increased use of adaptive designs in biomedical *and* psychosocial treatment (National Institute of Mental Health, 2008).

Despite the potential for adaptive strategies to yield improved outcomes and to enhance personalization of care, and despite the success of these strategies in the medication management of depression, few published studies provide data that directly inform adaptive approaches to psychotherapy for depression. Several studies, however, have tested adaptive approaches to psychotherapy for mixed psychiatric presentations. In an adaptive treatment for comorbid anxiety, depression and conduct problems in youth, therapists selected from treatment modules relevant to the most severe patient symptoms and were permitted to shift to alternative modules if other symptoms emerged as most salient (Weisz et al., 2012). The adaptive, module-based approach showed more rapid symptom improvement compared with standardized or procedurally specified treatments and treatment as usual. However, in that study, adaptation points were based on therapist judgment and were not operationalized, thus limiting dissemination and replication.

Lambert and colleagues developed a psychotherapy tracking system that uses early treatment response on a general symptom measure to derive a decision point regarding whether a patient will likely experience symptom improvement, no change in symptoms, or deterioration in symptoms (Hannan et al., 2005; Harmon, Hawkins, Lambert, Slade, & Whipple, 2005; M. Lambert et al., 2001). In heterogeneous samples of outpatient psychotherapy patients, providing clinicians with this feedback improved treatment outcomes compared with no clinician feedback; this finding was driven largely by a decreased proportion of patients who worsened over the course of psychotherapy in the feedback condition (M. J. Lambert, Harmon, Slade, Whipple, & Hawkins, 2005; M. J. Lambert & Shimokawa, 2011; Shimokawa, Lambert, & Smart, 2010). In an extension of this work, therapists were provided feedback along with clinical support tools (e.g. alliance repair strategies, social support strategies, motivation assessment) to help those at risk of worsening during treatment (Harmon et al., 2005). However, there was no additive advantage for feedback + clinical support compared to feedback alone (M. J. Lambert & Shimokawa, 2011). Thus, the limited data on adaptive strategies in psychotherapy are promising but not specific to depression. Furthermore, within depression treatment, remission of symptoms has been identified as a target treatment goal (Rush, Kraemer, et al., 2006). However, we know of no published research designed to establish decision points that could be used to signal a change of strategy with the goal of boosting remission rates in psychotherapy for depression.

As a first step toward developing adaptive treatment for depression, we aim to develop an operationalized strategy for identifying, as early in treatment as possible, those people who are unlikely to remit during the remainder of an ongoing course of psychotherapy for depression. This critical decision point would suggest *when* incorporating an alternative treatment strategy will likely be useful. In cognitive-behavioral therapy for depression, much symptomatic improvement occurs within the first 3–4 weeks of treatment (Ilardi & Craighead, 1994), and early treatment response robustly predicts remission in depression treatment (Tadic et al., 2010). This suggests early symptom change as a good candidate for a decision rule on whether someone is likely to remit. We therefore focus on identifying the earliest treatment time point at which percent symptom reduction from baseline is a clinically useful predictor of eventual remission as well as the level of improvement at that time point that optimally predicts remission. Using the large archival data set of a 12-week randomized controlled trial of treatment for chronic depression, we used receiver operating curve (ROC), a signal detection analysis, at various time points. Symptom reduction cut

points were identified in half the sample and then evaluated for predictive performance in the other half of the sample.

Methods

Design

This study is a secondary data analysis from the randomized acute treatment phase of a large trial for chronic depression (Keller et al., 2000). Participants were recruited from 12 sites and randomly assigned to receive 12 weeks of psychotherapy, antidepressant medication or their combination. The study was approved by institutional review boards at each of the recruiting sites and written informed consent was obtained from participants. Findings from the main outcome study showed that although participants improved substantially in all three treatment arms, those receiving combined treatment obtained significantly greater symptom reduction and significantly higher response and remission rates: Remission rates for the completer sample were 24% for psychotherapy only, 22% for medication only and 42% for combination treatment (Keller et al., 2000). Because the current analysis was designed to identify decision points for those receiving psychotherapy, analyses excluded participants assigned to the medication only treatment arm. Primary analyses for the current study were conducted in two phases: Phase I, in which triage cut points were identified and Phase II, in which identified cut points were tested in an independent sample, as recommended by Kraemer (1992). Analyses were conducted separately for the psychotherapy only and combination treatment arms.

Participants

Participants were outpatients aged 18–75 who met DSM-IV (American Psychiatric Association, 2000) criteria for current major depressive disorder (MDD) as determined by the Structured Clinical Interview for Axis I Disorders, Patient Edition (SCID-I/P) (First, Spitzer, Gibbon, & Williams, 1995). The parent study was focused on chronic forms of depression. Eligible participants had experienced a current MDD episode that lasted 2 or more years continuously or with incomplete interepisode recovery and/or an MDD episode superimposed on preexisting dysthymic disorder (“double depression”). In addition, participants scored at least 20 on the 24-item Hamilton Rating Scale for Depression (HRSD) (Hamilton, 1967) at screening and baseline. Additional details regarding exclusion criteria and the parent study sample appear in Keller et al. (2000). Because the remission definition required data from Weeks 10 and 12, the sample of the current study consisted only of those participants who completed 12 weeks of acute treatment.

Treatments

Psychotherapy consisted of Cognitive Behavioral Analysis System of Psychotherapy (CBASP), a time-limited psychotherapy designed specifically for people experiencing chronic forms of depression (McCullough, 2000). CBASP emphasizes elements of traditional cognitive-behavioral therapy while also incorporating aspects of interpersonal psychotherapy. A core procedure in CBASP is “situational analysis,” during which patients examine their thoughts and behaviors in the context of actual and desired outcomes of discrete interpersonal interactions. CBASP provides therapists with explicit procedures for highlighting incongruities between the therapeutic relationship and relationship expectations of patients. Following the CBASP manual (McCullough, 2000), CBASP sessions occurred twice weekly during weeks 1–4 and weekly during weeks 5–12. Patients having difficulty mastering situational analysis were offered twice weekly sessions through week 8. CBASP therapists were experienced: All had 2 or more years experience post Ph.D. or M.D. and 5 or more years experience post M.S.W. Therapists attended a two-day CBASP training and were certified based on videotaped sessions of pre-trial pilot cases. All study sessions were

also videotaped and supervisors reviewed videotapes weekly to assess adherence to the CBASP manual.

Those receiving combined treatment also received pharmacotherapy with the antidepressant nefazodone. Nefazodone was prescribed at an initial dose of 200 mg per day in week 1; increased to 300 mg per day during the second week. Thereafter, the dose was increased weekly in increments of 100 mg daily to a maximum of 600 mg per day until maximum efficacy and tolerability was achieved. Pharmacotherapy visits were 15–20 minutes in length, occurring weekly during weeks 1–4 and every other week thereafter. Following the Fawcett et al. (1987) clinical management manual, pharmacotherapy focused on discussion of concomitant medications and review of symptoms and side effects; formal psychotherapeutic intervention was proscribed (Fawcett, Epstein, Fiester, Elkin, & Autry, 1987).

Measures

Inventory of Depressive Symptoms-Self Report (IDS-SR)—Self-reported depressive symptoms were measured using the 30-item IDS-SR (Rush, Gullion, Basco, Jarrett, & Trivedi, 1996). Patients completed the IDS-SR at baseline, weekly from weeks 1–4, and bi-weekly from weeks 6–12. IDS-SR scores range from 0–84, with higher scores indicating greater depressive severity. The IDS-SR has shown good internal consistency in outpatients diagnosed with depression (Rush et al., 1996). It correlates with other self-report and clinician-rated measures of depressive symptoms and shows sensitivity to change in depressive symptoms that is comparable to clinician-ratings (Biggs et al., 2000; Rush, Trivedi, et al., 2006). As a self-report measure, the IDS-SR does not require a trained clinical rater, and is thus likely to facilitate a more feasible standardized data collection in real world settings as compared to a clinician-rated interview such as the HRSD. Therefore, for the current study, symptom change scores were derived from IDS-SR data and served as the predictor variable for end-of-treatment remission vs. non-remission as determined by the HRSD.

24-item Hamilton Rating Scale for Depression (HRSD)—Trained raters assessed depressive symptoms at baseline, weekly from weeks 1–4, and bi-weekly from weeks 6–12 using the 24-item HRSD (Hamilton, 1967). Raters were certified in administration of the HRSD and were blind to treatment condition. Remission during active treatment was defined *a priori* as a HRSD score of 8 or less at weeks 10 and 12.

Analytic Approach

The Phase I sample comprised approximately half of the cases, randomly selected from each treatment arm using the “select random cases” procedure in SPSS. The Phase II sample consisted of the remaining cases. Thus, Phase I and II samples were independent. The flow of participants from screening to the analyzable sample is depicted in Figure 1.

In Phase I, receiver operating curve (ROC) analyses (Kraemer, 1992) were conducted to explore whether percent symptom reduction on IDS-SR at weeks 4, 6, and 8 predicted remission status at week 12. In the absence of theoretical reasoning otherwise, we sought a cut point that weighted the value of sensitivity and specificity equally. We used percent symptom reduction rather than raw score change, in order to maximize the potential generalizability of identified cut points to other self-report measures of depressive symptoms. Weeks 4, 6, and 8 were selected as potentially clinically useful decision points because these time points represented a window around the midpoint of acute phase treatment and because prior research suggests that considerable symptomatic change takes place by week 4 (Ilardi & Craighead, 1994).

ROC analyses were conducted using the ROC4 program (available at <http://www.stanford.edu/~yesavage/ROC.html>). ROC is a non-parametric technique used to explore predictors of a dichotomous outcome variable. ROC is well suited to answering the question of interest as it can be used to identify clinically meaningful subgroups who may need tailoring of interventions (Kraemer, 1992). In ROC, when a variable is a significant predictor of the dichotomous outcome, the variable is assessed at all possible cut points to determine a cut point that maximizes sensitivity (i.e., likelihood of predicting true positives) and specificity (i.e. likelihood of predicting true negatives) according to a criterion designated by the user. In these analyses, the criterion was set at .50, giving equal weight to sensitivity and specificity in order to identify the cut point that would correctly classify the largest number of cases for eventual remission status. Phase I ROC analyses were conducted separately by treatment arm (combination or psychotherapy only) and separately for each treatment time point (week 4, 6, or 8), yielding a total of six separate ROC analyses. The outcome (criterion) variable of interest for each analysis was remission at the end of the acute treatment phase (remission vs. nonremission).

In Phase II, the performance of each cut point identified in Phase I was evaluated in the remainder of the sample. The rationale for using this independent sample for performance evaluation was twofold. First, ROC is an exploratory, hypothesis-generating approach that may capitalize on idiosyncrasies of the data (Kraemer, 1992). Thus, examining the performance of the cut points in a separate sample allows stronger inference regarding their generalizability. Second, the use of multiple ROC analyses in Phase I may have inflated the likelihood of Type I error. Thus, examination of the cut points in a separate sample served to check on the validity and potential clinical utility of the identified cut points. The following performance indicators of binary classification were calculated for identified cut points: Sensitivity, specificity, positive predictive value, negative predictive value and Cohen's kappa. McNemar's chi-square was used to compare classification accuracy between different assessment points; McNemar analyses were conducted separately for actual remitters and actual nonremitters as recommended by Trajman & Luiz (2008).

Results

Descriptive Characteristics

Remission rates did not differ significantly between the Phase I and Phase II samples for those receiving combination, $\chi^2 = .00$, $p = .99$, or those receiving psychotherapy only, $\chi^2 = .05$, $p = .83$. Descriptive characteristics of the sample appear in Table 1.

Phase I: Identification of Optimal Cut points

Results of Phase I ROC analyses appear in Figures 2a and 2b. Within the combined treatment group, analyses yielded the following statistics for predicting end of treatment HRSD status with percent IDS-SR symptom reduction: week 4: $\chi^2 = 6.26$, Cohen's kappa = .27; week 6: $\chi^2 = 16.17$, Cohen's kappa = .43; week 8: $\chi^2 = 33.56$, Cohen's kappa = .62. Within the psychotherapy only treatment, results were: week 4: $\chi^2 = 4.05$, Cohen's kappa = .22; week 6: $\chi^2 = 8.49$, Cohen's kappa = .35; week 8: $\chi^2 = 7.32$, Cohen's kappa = .29. Performance indicators for cut points are summarized in Table 2.

Phase II: Performance of Identified Cut points in an Independent Sample

Using the cut point criteria identified in Phase I, classification and performance indicators were calculated for Week 6 and Week 8 using the independent Phase II sample. Performance criteria appear in Tables 3a and 3b. Performance of Week 4 criteria were not calculated due to small effect sizes (Cohen's kappa's < .30) in Phase I ROCs. Within combination treatment, chi-square analyses showed no significant differences in sensitivity,

$\chi^2 = .00$, $p = 1.00$, or specificity, $\chi^2 = .44$, $p = .51$, between the cutoff values of weeks 6 and 8 and the week 6 Cohen's kappa of .42 was very similar to the week 8 value of .45. Within psychotherapy only treatment, the week 6 cut point was more sensitive than the week 8 cut point in predicting nonremission, $\chi^2 = 8.10$, $p < .01$, but was less specific, $\chi^2 = 5.14$, $p < .05$, with only 7 of 18 (39%) remitters being correctly classified at week 6. Reflecting this difference, the week 8 Cohen's kappa value of .49 was somewhat higher than the week 6 value of .41.

Discussion

This study identified critical decision points for identifying those people who are unlikely to show remission of chronic depression symptoms over the course of 12 weeks of structured psychotherapy, with or without antidepressant medication. Although symptom reduction before the midpoint of treatment did not meaningfully predict remission status, nonremission was detectable as early as the mid time-point of treatment (week 6 in this study). Mid time-point prediction was particularly good among those receiving psychotherapy combined with antidepressant medication. However, among those receiving psychotherapy only, there was a relatively high false positive rate in predicting nonremission at treatment midpoint that was improved when using a later time point (week 8). The finding that prognosis can be established more quickly among those receiving combination treatment compared to psychotherapy alone is consistent with a prior analysis of the same data set (Manber et al., 2008) and with a prior adolescent finding showing that time to stable response is more rapid in combination treatment (Kratohvil et al., 2006). Thus, among those receiving combination treatment, approximately 6 weeks of treatment is likely to be a reliable time for predicting nonremission based on treatment response. In addition, when rapid detection of likely nonremission is of primary importance (e.g. session number is limited by insurance reimbursement), 6 weeks may also be a useful time point for identifying likely nonremitters in psychotherapy only. However, using a slightly later time point is likely to minimize false identification of nonremission.

This study highlights that those unlikely to remit can be identified prior to treatment termination. Doing so may enable a change in treatment approach and improved overall outcomes. Although identifying critical decision points is a first step in developing adaptive interventions, the identified decision points are exploratory and preliminary. It will be important to examine them prospectively in future research in order to examine their relationship to symptom outcome and attrition. Another important next step in designing adaptive psychotherapy interventions is to identify what such adaptations might entail. Although the decision points identified in this study offers guidelines on *when* to adapt, additional research is needed to investigate *what* adaptations are most helpful. In addition, although the current study focused on non-remitters, an alternative direction for future adaptive design research in depression might focus on optimal and efficient course of maintenance treatment among those who do remit during acute phase treatment. For example, adaptive designs could be used to identify decision points regarding the frequency of "booster" psychotherapy sessions and/or continuation versus discontinuation of antidepressants. This application of smart designs has been used in a recent study examining treatment for pediatric and adolescent anxiety (Almirall, Compton, Rynn, Walkup, & Murphy, 2012).

Several limitations warrant consideration in evaluating these preliminary findings. First, our sample comprised only people meeting criteria for chronic depression who enrolled in a time-limited and highly structured clinical trial, which may limit the generalizability of the findings. Second, because psychotherapy sessions took place twice weekly during the first month of treatment, critical decision time points in this study may have been accelerated

compared to treatments that include once weekly sessions throughout. In future research, it may be useful to identify critical decision points based on number of sessions, which would increase applicability to less structured clinical practice settings where the frequency of sessions may vary. Third, because our *a priori* definition of remission required data from weeks 10 and 12, we included only those participants who completed acute phase treatment in the current analysis. Predictive cutoffs may be different for an intent-to-treat sample, depending on reasons for non-completion. For example, a lower symptom reduction cut point would have been identified with inclusion of those who responded particularly well to treatment and were less likely to attend later sessions once they were feeling better. Conversely, a higher symptom reduction cut point would have been identified if those who responded particularly poorly to treatment were more likely to drop out before the end of treatment. In future research, using alternative definitions of remission and/or data imputation techniques would enable inclusion of treatment noncompleters, thus allowing for formal testing of hypotheses regarding the effect of noncompleters on identified cut points.”

We also note that sensitivity and specificity of a criterion may vary when using the test in populations with different rates of remission (Kraemer, 1992). Within the psychotherapy only condition, the observed remission rate of 24% is within the range of remission observed in other studies of evidence-based psychotherapies for depression (Arnow & Hill, 2008), thus somewhat increasing confidence in the generalizability of the current results to other forms of depression. However, additional studies will be important to further assess generalizability to other forms of psychotherapy, psychotherapy of differing time lengths, and other delivery settings. Within the combination treatment condition, the 42% remission rate observed in the current sample resembles other samples of chronically depressed participants but is lower than that observed in trials that have enrolled participants with more acute or episodic depressions (Thase et al., 1997). Replicating these findings for acute depression may be needed before applying the cutoff more broadly in combination treatment. In general, future research is needed to evaluate how well the current findings will generalize to other depressed populations, treatment settings, and treatment lengths/types.

These limitations notwithstanding, the current study has a number of strengths. The data were taken from a large and rigorously evaluated sample, treatments included psychotherapy only and combination treatment arms, and trained raters administered the outcome measure specifying remission (HRSD). The large sample size allowed for the split-sample methodology, strengthening inference regarding the identified cut point criteria. Basing cut points on a different symptom measure (IDS-SR) than was used to define gold standard remission (HRSD) provides convergent evidence that the symptom reduction cuts signify symptom improvement, and reduces the possibility that findings are an artifact of overlap between the definition of remission and identified cut points. Furthermore, basing the clinical decision points on a self-report measure provides a strategy that can be feasibly incorporated into future studies or clinical practice.

We believe these findings are a useful initial step in developing adaptive psychotherapy approaches for depression. In addition, we hope these findings may spur further research using this analytic approach with different treatment types and lengths, and different populations. Toward that goal, we wish to point out several methodological choices that could be modified in future research that uses this analytic approach. First, in the absence of theoretical rationale otherwise, we elected to run the analysis such that sensitivity and specificity were given equal importance (i.e. to identify a cut point that would correctly classify the greatest number of people). However, depending on the treatment situation and clinician preference, guidelines that place a greater emphasis on specificity (i.e. minimizing false positives) or a greater emphasis on sensitivity (minimizing false negatives) would be

useful. For example, a clinician may prefer to use a guideline with greater specificity, in essence, to “stay the course” with a given treatment unless it is particularly unlikely to be effective. The ROC4 program allows the flexibility to adjust for preferred sensitivity and specificity weighting. Indeed, one prior study has used such an approach in identifying predictors of response to pharmacotherapy in late-life depression (Andreescu et al., 2008). Second, because of the relative dearth of information regarding psychotherapy clinical decision points, we focused the current study on psychotherapy treatments. However, this approach to identifying decision points could also be applied to biomedical treatments, including pharmacotherapy, and other forms of treatment. Third, we chose to use the decision point metric of percent symptom reduction, rather than raw score change on a symptom measure. Our decision was based on the idea that percent symptom reduction more easily allows comparison with other studies that may use a different symptom measure, and to enhance interpretability for those readers less familiar with the IDS-SR, the symptom measure used in this study. However, in future research, absolute score change could be used in lieu of, or in addition to percent symptom reduction in identifying decision cut points. Absolute change scores may be preferable for studies using symptom measures in widespread clinical use (e.g. The Beck Depression Inventory).

Conclusion

Those unlikely to remit from chronic depression over the course of 12 weeks of structured psychotherapy can be identified prior to end of treatment based on percent symptom reduction from baseline. Symptom reduction was meaningfully predictive as early as week 6, particularly among those receiving psychotherapy in combination with antidepressant medication. However, week 8 may be a more useful critical decision point among those receiving psychotherapy only. Future research should clarify the generalizability of the decision points identified in this study and apply this approach to additional treatment types, lengths and populations. However, the current findings provide initial evidence that cutoffs on the simple and feasible metric of self-reported depressive symptom reduction can provide an operationalized decision point for modifying the treatment approach during psychotherapy for depression.

Acknowledgments

Financial support for data collection was provided by Bristol Myers Squibb.

This work was supported in part by the National Institute of Mental Health T32-MH019938-18 awarded to Alan F. Schatzberg. In addition, the authors wish to thank the many study personnel who assisted in study execution and data collection as well as the study participants.

References

- Almirall D, Compton SN, Gunlicks-Stoessel M, Duan N, Murphy SA. Designing a pilot sequential multiple assignment randomized trial for developing an adaptive treatment strategy. *Stat Med*. 2012;10.1002/sim.4512
- Almirall D, Compton SN, Rynn MA, Walkup JT, Murphy SA. SMARTer discontinuation trial designs for developing an adaptive treatment strategy. *J Child Adolesc Psychopharmacol*. 2012; 22(5):364–374.10.1089/cap.2011.0073 [PubMed: 23083023]
- American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 4. 2000. Text Revision
- Andreescu C, Mulsant BH, Houck PR, Whyte EM, Mazumdar S, Dombrovski AY, Reynolds CF 3rd. Empirically derived decision trees for the treatment of late-life depression. *Am J Psychiatry*. 2008; 165(7):855–862.10.1176/appi.ajp.2008.07081340 [PubMed: 18450930]

- Arnow BA, Hill K. The efficacy of CBT, IPT, and related psychotherapies for depression: A review by treatment phase. *Depression: Mind and Body*. 2008; 3(4):15–27.
- Biggs MM, Shores-Wilson K, Rush AJ, Carmody TJ, Trivedi MH, Crismon ML, Mason M. A comparison of alternative assessments of depressive symptom severity: a pilot study. *Psychiatry Res*. 2000; 96(3):269–279. [PubMed: 11084222]
- Crismon ML, Trivedi M, Pigott TA, Rush AJ, Hirschfeld RM, Kahn DA, Thase ME. The Texas Medication Algorithm Project: report of the Texas Consensus Conference Panel on Medication Treatment of Major Depressive Disorder. *J Clin Psychiatry*. 1999; 60(3):142–156. [PubMed: 10192589]
- Cuijpers P, Andersson G, Donker T, van Straten A. Psychological treatment of depression: results of a series of meta-analyses. *Nord J Psychiatry*. 2011; 65(6):354–364.10.3109/08039488.2011.596570 [PubMed: 21770842]
- Cuijpers P, van Straten A, Warmerdam L, Andersson G. Psychotherapy versus the combination of psychotherapy and pharmacotherapy in the treatment of depression: a meta-analysis. *Depress Anxiety*. 2009; 26(3):279–288.10.1002/da.20519 [PubMed: 19031487]
- Fava M, Davidson KG. Definition and epidemiology of treatment-resistant depression. *Psychiatr Clin North Am*. 1996; 19(2):179–200. [PubMed: 8827185]
- Fava M, Rush AJ, Trivedi MH, Nierenberg AA, Thase ME, Sackeim HA, Kupfer DJ. Background and rationale for the sequenced treatment alternatives to relieve depression (STAR*D) study. *Psychiatr Clin North Am*. 2003; 26(2):457–494. x. [PubMed: 12778843]
- Fawcett J, Epstein P, Fiester SJ, Elkin I, Autry JH. Clinical management--imipramine/placebo administration manual. NIMH Treatment of Depression Collaborative Research Program. *Psychopharmacol Bull*. 1987; 23(2):309–324. [PubMed: 3303100]
- First, MB.; Spitzer, RL.; Gibbon, M.; Williams, JBW. Structured Clinical Interview for DSM-IV Axis-I Disorders, Patient Edition (SCID-I/P, version 2.0). New York: Bio-metrics Research Dept., New York State Psychiatric Institute; 1995.
- Hamilton M. Development of a rating scale for primary depressive illness. *Br J Soc Clin Psychol*. 1967; 6(4):278–296. [PubMed: 6080235]
- Hannan C, Lambert MJ, Harmon C, Nielsen SL, Smart DW, Shimokawa K, Sutton SW. A lab test and algorithms for identifying clients at risk for treatment failure. *J Clin Psychol*. 2005; 61(2):155–163.10.1002/jclp.20108 [PubMed: 15609357]
- Harmon C, Hawkins EJ, Lambert MJ, Slade K, Whipple JS. Improving outcomes for poorly responding clients: the use of clinical support tools and feedback to clients. *J Clin Psychol*. 2005; 61(2):175–185.10.1002/jclp.20109 [PubMed: 15609355]
- Ilardi SS, Craighead WE. The Role of Nonspecific Factors in Cognitive-Behavior Therapy for Depression. *Clinical Psychology*. 1994; 1(2):138–156.
- Keller MB, McCullough JP, Klein DN, Arnow B, Dunner DL, Gelenberg AJ, Zajecka J. 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):1462–1470.10.1056/nejm200005183422001 [PubMed: 10816183]
- Kraemer, HC. Evaluating medical tests: Objective and quantitative guidelines. Newbury Park, CA: Sage; 1992.
- Kratochvil C, Emslie G, Silva S, McNulty S, Walkup J, Curry J, March J. Acute time to response in the Treatment for Adolescents with Depression Study (TADS). *J Am Acad Child Adolesc Psychiatry*. 2006; 45(12):1412–1418.10.1097/01.chi.0000237710.73755.14 [PubMed: 17135986]
- Lambert M, Whipple J, Smart D, Vermeersch D, Nielsen S, Hawkins E. The effects of providing therapists with feedback on patient progress during psychotherapy: Are outcomes enhanced? *Psychotherapy Research*. 2001; 11(1):49–68.
- Lambert MJ, Harmon C, Slade K, Whipple JL, Hawkins EJ. Providing feedback to psychotherapists on their patients' progress: clinical results and practice suggestions. *J Clin Psychol*. 2005; 61(2):165–174.10.1002/jclp.20113 [PubMed: 15609358]
- Lambert MJ, Shimokawa K. Collecting client feedback. *Psychotherapy (Chic)*. 2011; 48(1):72–79.10.1037/a0022238 [PubMed: 21401277]

- Lavori PW, Dawson R. A design for testing clinical strategies: Biased adaptive within-subject randomization. *Journal of Royal Statistical Society Series A: Statistics in Society*. 2000; (163):29–38.
- Lavori PW, Dawson R, Rush AJ. Flexible treatment strategies in chronic disease: clinical and research implications. *Biol Psychiatry*. 2000; 48(6):605–614. [PubMed: 11018231]
- Manber R, Kraemer H, Arnow B, Trivedi M, Rush A, Thase M, Keller M. Faster remission of Chronic Depression with combined psychotherapy and medication than with either therapy alone. *Journal of Consulting and Clinical Psychology*. 2008; 26(3):459–467. [PubMed: 18540739]
- McCullough, JP. *Treatment for Chronic Depression: Cognitive Behavioral Analysis System of Psychotherapy*. New York, NY: Guilford Press; 2000.
- Murphy S. An experimental design for development of adaptive treatment strategies. *Stat Med*. 2005; 24:1455–1481. [PubMed: 15586395]
- Murphy SA, Oslin DW, Rush AJ, Zhu J. Methodological challenges in constructing effective treatment sequences for chronic psychiatric disorders. *Neuropsychopharmacology*. 2007; 32(2):257–262.10.1038/sj.npp.1301241 [PubMed: 17091129]
- National Institute of Mental Health. NIMH Strategic Plan. 2008. Retrieved from <http://www.nimh.nih.gov/about/strategic-planning-reports/nimh-strategic-plan-2008.pdf>
- Pigott HE, Leventhal AM, Alter GS, Boren JJ. Efficacy and effectiveness of antidepressants: current status of research. *Psychother Psychosom*. 2010; 79(5):267–279.10.1159/000318293 [PubMed: 20616621]
- Quitkin FM, McGrath PJ, Stewart JW, Ocepek-Welikson K, Taylor BP, Nunes E, Klein DF. Chronological milestones to guide drug change. When should clinicians switch antidepressants? *Arch Gen Psychiatry*. 1996; 53(9):785–792. [PubMed: 8792755]
- Richards D. Prevalence and clinical course of depression: a review. *Clin Psychol Rev*. 2011; 31(7): 1117–1125.10.1016/j.cpr.2011.07.004 [PubMed: 21820991]
- Rush AJ, Fava M, Wisniewski SR, Lavori PW, Trivedi MH, Sackeim HA, Niederehe G. Sequenced treatment alternatives to relieve depression (STAR*D): rationale and design. *Control Clin Trials*. 2004; 25(1):119–142. [PubMed: 15061154]
- Rush AJ, Gullion CM, Basco MR, Jarrett RB, Trivedi MH. The Inventory of Depressive Symptomatology (IDS): psychometric properties. *Psychol Med*. 1996; 26(3):477–486. [PubMed: 8733206]
- Rush AJ, Kraemer HC, Sackeim HA, Fava M, Trivedi MH, Frank E, Schatzberg AF. Report by the ACNP Task Force on response and remission in major depressive disorder. *Neuropsychopharmacology*. 2006; 31(9):1841–1853.10.1038/sj.npp.1301131 [PubMed: 16794566]
- Rush AJ, Trivedi MH, Wisniewski SR, Nierenberg AA, Stewart JW, Warden D, Fava M. Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D report. *Am J Psychiatry*. 2006; 163(11):1905–1917.10.1176/appi.ajp.163.11.1905 [PubMed: 17074942]
- Shimokawa K, Lambert MJ, Smart DW. Enhancing treatment outcome of patients at risk of treatment failure: meta-analytic and mega-analytic review of a psychotherapy quality assurance system. *J Consult Clin Psychol*. 2010; 78(3):298–311.10.1037/a0019247 [PubMed: 20515206]
- Spielmans GI, Berman MI, Usitalo AN. Psychotherapy versus second-generation antidepressants in the treatment of depression: a meta-analysis. *J Nerv Ment Dis*. 2011; 199(3):142–149.10.1097/NMD.0b013e31820caefb [PubMed: 21346483]
- Tadic A, Helmreich I, Mergl R, Hautzinger M, Kohnen R, Henkel V, Hegerl U. Early improvement is a predictor of treatment outcome in patients with mild major, minor or subsyndromal depression. *J Affect Disord*. 2010; 120(1–3):86–93.10.1016/j.jad.2009.04.014 [PubMed: 19428118]
- Thase ME, Greenhouse JB, Frank E, Reynolds CF 3rd, Pilskonis PA, Hurley K, Kupfer DJ. Treatment of major depression with psychotherapy or psychotherapy-pharmacotherapy combinations. *Arch Gen Psychiatry*. 1997; 54(11):1009–1015. [PubMed: 9366657]
- Trajman A, Luiz RR. McNemar chi2 test revisited: comparing sensitivity and specificity of diagnostic examinations. *Scand J Clin Lab Invest*. 2008; 68(1):77–80.10.1080/00365510701666031 [PubMed: 18224558]

- Trivedi MH, Rush AJ, Crismon ML, Kashner TM, Toprac MG, Carmody TJ, Shon SP. Clinical results for patients with major depressive disorder in the Texas Medication Algorithm Project. *Arch Gen Psychiatry*. 2004; 61(7):669–680.10.1001/archpsyc.61.7.669 [PubMed: 15237079]
- Weisz JR, Chorpita BF, Palinkas LA, Schoenwald SK, Miranda J, Bearman SK, Gibbons RD. Testing standard and modular designs for psychotherapy treating depression, anxiety, and conduct problems in youth: a randomized effectiveness trial. *Arch Gen Psychiatry*. 2012; 69(3):274–282.10.1001/archgenpsychiatry.2011.147 [PubMed: 22065252]

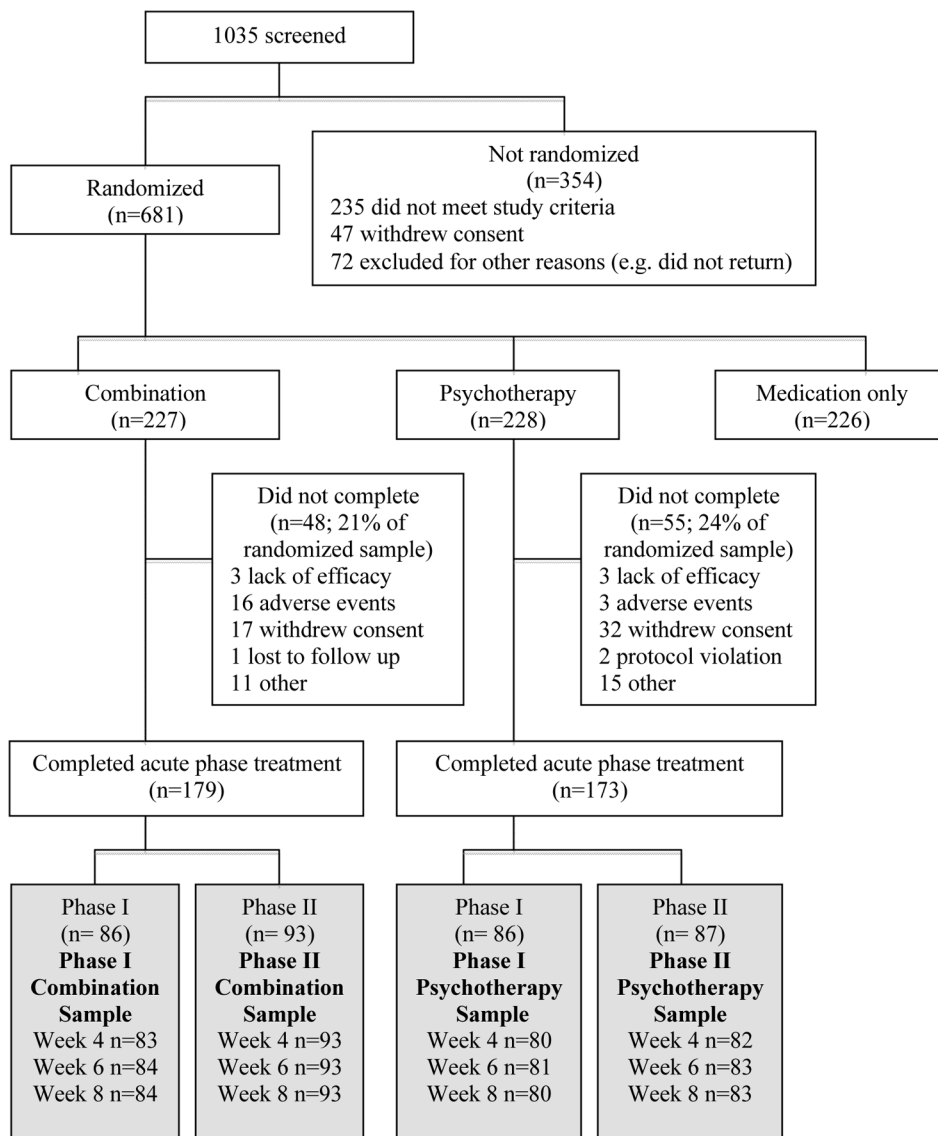


Figure 1.
Participant flow

Figure 2a:

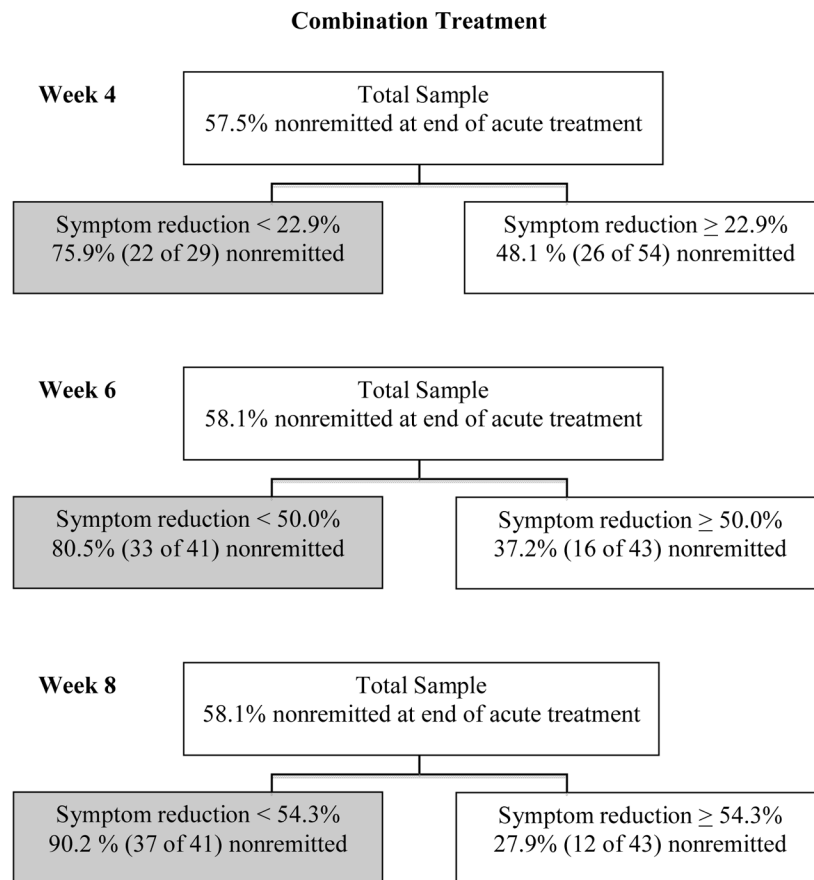


Figure 2b:

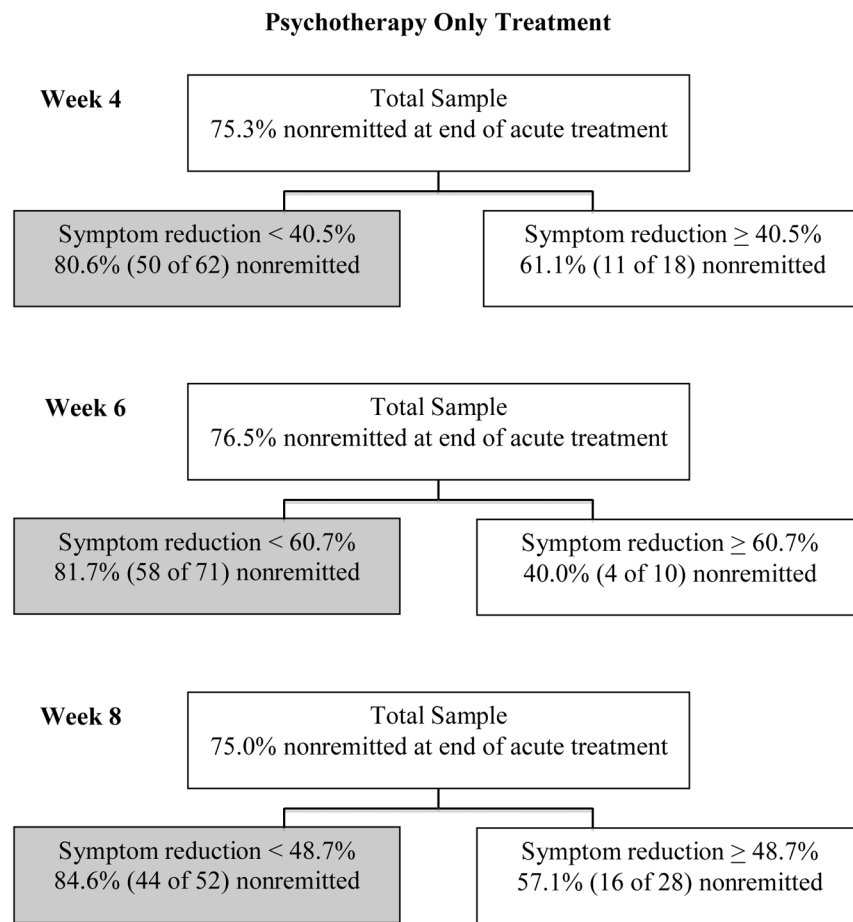


Figure 2.

Figure 2a: Phase I ROC Analyses for Combination Treatment

Figure 2b: Phase I ROC Analyses for Psychotherapy Only Treatment

Table 1

Descriptive characteristics of the sample

	Combination		Psychotherapy		Overall
	Phase I	Phase II	Phase I	Phase II	
n	86	93	86	87	352
Sex, % female	73.3	63.4	70.9	55.2	65.6
Race					
White	91.9	95.7	90.7	90.8	92.3
Black	4.7	1.1	5.8	2.3	3.4
Asian	0	0	1.2	3.4	1.1
Hispanic	2.3	0	1.2	1.1	1.1
Other	1.2	3.2	1.2	2.3	2.0
Age in years, <i>M(SD)</i>	44.9 (9.8)	44.7 (9.9)	44.9 (10.0)	42.9 (11.0)	44.3 (10.2)
Marital status, %					
Married or cohabiting	43.0	49.5	38.4	48.3	44.9
Single	27.9	18.3	30.2	33.3	27.3
Widowed	2.3	1.1	1.2	1.1	1.4
Divorced or separated	26.7	31.2	30.2	17.2	26.4
Depression diagnosis, %					
Chronic major depression	31.4	34.4	39.5	31.0	34.1
MDD with dysthymic disorder	41.9	39.8	36.0	46.0	40.9
Recurrent, with incomplete interepisode remission	26.7	25.8	24.4	23.0	25.0
Baseline HRSD score, <i>M(SD)</i>	27.6 (5.1)	26.9 (4.9)	26.7 (4.5)	26.2 (5.3)	26.9 (4.9)
Baseline IDS-SR score, <i>M(SD)</i>	40.9 (8.8)	38.8 (8.3)	38.7 (8.8)	39.3 (7.9)	39.4 (8.4)
Age at MDD onset, <i>M(SD)</i>	29.1 (13.7)	26.7 (13.1)	26.9 (11.8)	27.9 (14.2)	27.6 (13.2)
Duration of current episode in years, <i>M(SD)</i>	9.0 (10.1)	7.8 (9.6)	9.6 (10.9)	7.4 (10.4)	8.5 (10.2)
History of anxiety disorder, %	34.9	33.3	25.6	26.4	30.1
Prior treatment with psychotherapy and antidepressants, %	52.3	51.6	39.5	47.1	47.8
Prior psychotherapy, %	69.8	69.9	61.6	64.4	66.5
No prior depression treatment, %	15.1	14.0	19.8	24.1	18.2
Baseline Global Assessment of Functioning Score, <i>M(SD)</i>	52.9 (5.5)	54.4 (5.7)	53.9 (5.1)	53.7 (6.0)	53.7 (5.6)

Table 2
Performance of ROC Symptom Reduction Criteria In Predicting Depression Nonremission in Phase I

	Combination treatment				Psychotherapy only					
	Value	95% CI Lower	95% CI Upper	Value	95% CI Lower	95% CI Upper	Value	95% CI Lower	95% CI Upper	
Week 4										
Specificity	.80	.63	.91	.37	.17	.61				
Sensitivity	.46	.32	.61	.82	.70	.90				
Negative predictive value	.52	.38	.65	.39	.18	.64				
Positive predictive value	.76	.56	.89	.81	.68	.89				
Cohen's kappa	.24	.06	.43	.19	-.05	.43				
Week 6										
Specificity	.77	.59	.89	.32	.14	.57				
Sensitivity	.67	.52	.80	.93	.84	.98				
Negative predictive value	.63	.47	.77	.60	.27	.86				
Positive predictive value	.80	.65	.91	.82	.70	.90				
Cohen's kappa	.43	.24	.62	.30	.06	.55				
Week 8										
Specificity	.89	.72	.96	.60	.36	.80				
Sensitivity	.76	.61	.86	.73	.60	.84				
Negative predictive value	.72	.56	.84	.42	.25	.63				
Positive predictive value	.90	.76	.97	.85	.71	.93				
Cohen's kappa	.62	.46	.79	.29	.08	.51				

Table 3a
 Classification of Actual Remitters and Nonremitters in Phase II Using Symptom Reduction Cut Points Identified in ROC Analyses

	Combination Treatment			Psychotherapy Only		
	Actual remitter	Actual nonremitter	Total	Actual remitter	Actual nonremitter	Total
Week 6						
Likely remitter	24	11	35	7	3	10
Likely nonremitter	14	41	55	11	62	73
Total	38	52	90	18	65	83
Week 8						
Likely remitter	26	12	38	14	13	27
Likely nonremitter	11	40	51	4	52	56
Total	37	52	89	18	65	83

Table 3b
 Performance of ROC Symptom Reduction Criteria In Predicting Depression Nonremission in Phase II

	Combination treatment				Psychotherapy only					
	Value	95% CI Lower	95% CI Upper	Value	95% CI Lower	95% CI Upper	Value	95% CI Lower	95% CI Upper	
Week 6										
Specificity	.62	.45	.76	.39	.18	.64	.64	.18	.64	
Sensitivity	.80	.66	.90	.95	.86	.99	.99	.86	.99	
Negative predictive value	.69	.51	.83	.70	.35	.92	.92	.35	.92	
Positive predictive value	.74	.61	.84	.85	.74	.92	.92	.74	.92	
Cohen's kappa	.42	.23	.61	.41	.16	.66	.66	.16	.66	
Week 8										
Specificity	.67	.50	.80	.78	.52	.93	.93	.52	.93	
Sensitivity	.78	.64	.88	.80	.68	.89	.89	.68	.89	
Negative predictive value	.68	.51	.82	.52	.32	.71	.71	.32	.71	
Positive predictive value	.76	.63	.86	.93	.82	.98	.98	.82	.98	
Cohen's kappa	.45	.26	.63	.49	.29	.69	.69	.29	.69	