



Published in final edited form as:

J Consult Clin Psychol. 2009 August ; 77(4): 775–787. doi:10.1037/a0015401.

Prediction of Response to Medication and Cognitive Therapy in the Treatment of Moderate to Severe Depression

Jay C. Fournier,

Department of Psychology, University of Pennsylvania

Robert J. DeRubeis,

Department of Psychology, University of Pennsylvania

Richard C. Shelton,

Department of Psychiatry, Vanderbilt University Medical Center

Steven D. Hollon,

Department of Psychology, Vanderbilt University

Jay D. Amsterdam, and

Department of Psychiatry, University of Pennsylvania School of Medicine

Robert Gallop

Department of Mathematics and Applied Statistics, West Chester University

Abstract

A recent randomized controlled trial found nearly equivalent response rates for antidepressant medications and cognitive therapy in a sample of moderate-to-severely depressed outpatients. In this article, we seek to identify the variables that were associated with response across both treatments as well as variables that predicted superior response in one treatment over the other. The sample consisted of 180 depressed outpatients: 60 of whom were randomly assigned to cognitive therapy; 120 were assigned to antidepressant medications. Treatment was provided for 16 weeks. Chronic depression, older age, and lower intelligence each predicted relatively poor response across both treatments. Three prescriptive variables were identified: marriage, unemployment, and having experienced a greater number of recent life events predicted superior response to cognitive therapy compared to antidepressant medications. Thus, six markers of treatment outcome were identified, each of which might be expected to carry considerable clinical utility. The three prognostic variables identify subgroups that might benefit from alternative treatment strategies; the three prescriptive variables identify groups who appear to respond particularly well to cognitive therapy.

Keywords

Treatment of Depression; Cognitive Therapy; Antidepressant Medications; Moderators; Prediction of Outcome

Correspondence concerning this article should be addressed to Jay C. Fournier, Department of Psychology, University of Pennsylvania, 3720 Walnut Street, Philadelphia, PA, 19104. jcf@sas.upenn.edu.

The data presented in this paper come from a clinical trial that was conducted between 1998 and 2003. The findings described in this manuscript have not been presented elsewhere in printed form, however, a subset of the findings were presented at the 2006 international meeting of the Society for Psychotherapy Research, Edinburgh, Scotland.

No authors have relevant conflicts of interest to disclose.

Cognitive therapy and pharmacotherapy are two well-validated, efficacious treatments for depression. Although the two treatments appear to be approximately equivalent in alleviating the acute symptoms of depression when large, aggregate treatment samples are compared (DeRubeis, Gelfand, Tang, & Simons, 1999), evidence is mounting to suggest that the efficacy of these treatments may be influenced to some degree by the traits and characteristics of the individuals presenting for treatment (Fournier et al., 2008; Hollon, Jarrett et al., 2005; Leykin et al., 2007; Rude & Rehm, 1991; Sotsky et al., 1991). In this article we attempt to further identify and examine such characteristics as we report analyses examining the prediction of treatment response in a recent randomized-controlled trial comparing cognitive therapy and antidepressant medication for the treatment of moderate to severe depression (DeRubeis et al., 2005; Hollon, DeRubeis et al., 2005).

There are two distinct ways in which a pre-treatment variable can participate in a predictive relationship with outcome. A *prognostic* variable is one that predicts outcome irrespective of the treatment, whereas a *prescriptive* variable (in the methodological literature this is often referred to as a moderator; Kraemer, Wilson, Fairburn, & Agras, 2002) predicts a different pattern of outcomes between two, or more, treatment modalities (Hollon & Beck, 1986). Prognostic variables can indicate which kinds of patients are especially refractory to treatment irrespective of the type of intervention. These variables may indicate a more severe subtype of depression, and their clinical utility derives from their ability to indicate which patients may require a more intensive course of treatment or a different modality of treatment. On the other hand, when two treatments are found not to differ in outcome on average, as in the present study, it may nonetheless be the case that pretreatment markers will be found that indicate a group of patients that fare considerably better in one treatment than the other. Such markers could be used prescriptively and would have a great deal of clinical utility, as they would allow individual patients to receive the treatment that is the most likely to lead to a rapid reduction in their depressive symptoms.

A complete review of the literature examining predictors of response to treatments for depression is beyond the scope of this article, as much of this literature has addressed the prediction of response within a single treatment modality (e.g., Hirschfeld et al., 1998; Szadoczky, Rozsa, Zambori, & Furedi, 2004; Thase et al., 1994; Trivedi, Morris, Grannemann, & Mahadi, 2005). For some baseline characteristics, relatively consistent findings have been reported across such studies, whereas inconsistent findings have been obtained with others. For example, in the treatment of depression with antidepressant medications, most investigators (Hirschfeld et al., 1998; Petersen et al., 2002; Sotsky et al., 1991; Szadoczky et al., 2004) have found that participant age has no effect on outcome; however, regarding gender, some find no effect (Hirschfeld et al., 1998; Petersen et al., 2002; Sotsky et al., 1991; Szadoczky et al., 2004), others find that males fare better (Mazure, Bruce, Maciejewski, & Jacobs, 2000), whereas still others find that females, on average, evidence a superior response (Trivedi et al., 2006). The results of such single-treatment studies may provide some limited information about prognosis for the specific treatment in question; however, they offer little guidance as to which treatment modality would be the most effective for a patient with the given characteristic. That is, even if such a single-treatment study were to find that older patients, for example, fared worse than younger patients in trials of cognitive therapy, this would provide at best limited guidance as to which is the optimal treatment for older patients. It is possible that the disparity between young and old might be even greater in other treatment modalities, such that cognitive therapy might still provide the best chance for successful outcome for older individuals, even though few such patients would be expected to respond.

There have been relatively few reports of predictive analyses from studies that have compared two or more treatments for depression. Perhaps the most widely known comparative trial is the National Institute of Mental Health's Treatment of Depression Collaborative Research

Project (TDCRP; Elkin et al., 1995). This study has provided researchers with a rich dataset for examining predictors of response in pharmacotherapy and cognitive therapy treatments. Although in the original analyses no overall differences between the two treatments were reported, secondary analyses revealed that cognitive therapy was less effective than medications, and no more effective than placebo, among their more severely depressed patients (Elkin et al., 1995). Had this finding been replicated, it would have provided valuable clinical information about the treatment of choice for patients with more severe depressions. As Jacobson and Hollon (1996) point out, however, this pattern was not robust across sites within the TDCRP sample. Furthermore, a subsequent mega-analysis (DeRubeis et al., 1999) that combined data from the TDCRP with data from three other randomized clinical trials found that cognitive therapy was at least as effective as antidepressant medication in the treatment of more-severely depressed patients. Indeed, the randomized-controlled trial from which the present study's data originate was initiated to compare response to cognitive therapy and antidepressant medications for more severely depressed patients. It found no evidence that depression severity was associated with differences in efficacy between the two treatments.

Melancholia is another feature of depression that is widely believed to be associated with superior response to medications over psychotherapy (Hollon, Jarrett et al., 2005). Indeed, the classification of melancholic depression was originally intended to identify those individuals who are particularly likely to respond to medications (Nelson, Mazure, & Jatlow, 1990). There is conflicting evidence, however, with some finding that it predicts poor response to antidepressants (Fava et al., 1997), some finding that it predicts favorable response (Heiligenstein, Tollefson, & Faries, 1994), and some finding no relationship with outcome (Nelson et al., 1990; Sotsky et al., 1991). Differences have emerged between the various antidepressant drug classes (see Esposito & Goodnick, 2003; Goodwin, 1993 for reviews), however, there is little evidence that melancholia, or any other variable that refers to a class or subtype of depression, is associated with differential response to medications versus cognitive therapy (Hollon, Jarrett et al., 2005; Jarrett et al., 1999).

Researchers have used the TDCRP dataset to explore other prescriptive factors that might differentiate response to cognitive therapy and antidepressant medications (Agosti & Ocepek-Welikson, 1997; Blatt, Quinlan, Pilkonis, & Shea, 1995; Blatt, Zuroff, Bondi, Sanislow, & Pilkonis, 1998; Sotsky et al., 1991), as well as data from other comparative treatment outcome studies (Blackburn & Moore, 1997; Jarrett et al., 1999). Although several history of illness, demographic, cognitive, and baseline functioning variables have been examined in these analyses, no prescriptive factors have been successfully identified. Part of the difficulty in uncovering these effects likely stems from the relatively low statistical power associated with the predictor-by-treatment-group interaction effects from ANCOVA, Moderated Multiple Regression, and Logistic Regression Models (Aguinis & Stone-Romero, 1997; Aiken & West, 1991), which have been the standard statistical approaches used to identify such prescriptive variables. With the advent of more sophisticated multilevel modeling techniques, researchers now have a more powerful set of statistical tools that make fewer statistical assumptions, produce more precise parameter estimates, and take full advantage of all data collected at each time point in clinical trials (Raudenbush & Bryk, 2002; Singer & Willett, 2003). In the present study, we make use of these techniques in an attempt to identify variables that could be used either prescriptively or prognostically to predict outcome in cognitive therapy and antidepressant medication treatment. We aim to balance the goal of replicating the efforts of previous research with a comprehensive exploration of variables that might predict outcome in the current dataset while at the same time attempting to balance the risks of type-I and type-II errors.

The main treatment outcome findings from the data used in this study have been reported elsewhere (DeRubeis et al., 2005; Hollon, DeRubeis et al., 2005). In addition, two planned

prediction analyses have been carried out using these data and have successfully identified prescriptive variables that moderate treatment response. Specifically, the pattern of results from one set of analyses indicated that patients with comorbid personality disorders were more likely to respond to medication treatment than to cognitive therapy, whereas the reverse pattern was found for patients without comorbid personality disorders – namely that these patients responded better to cognitive therapy than to medications (Fournier et al., 2008). The second set of analyses indicated that patients who had undertaken previous attempts at antidepressant medication were less likely to respond to medication than to cognitive therapy, whereas no differences in outcome between the two treatments were found for medication-naïve patients (Leykin et al., 2007).

In the present investigation we explored the prognostic and prescriptive predictive value of variables that were assessed at pre-treatment. Because the number of candidate variables was large, we developed a data-analytic approach that maximizes the chances of identifying important prognostic and prescriptive markers of response without increasing unduly the likelihood of capitalizing on chance. After we identified prognostic and prescriptive variables, using this method, we examined whether each of those that were identified remained a significant predictor of outcome when all of the predictors were tested simultaneously in the same statistical model. In addition, we attempted to identify predictors of attrition, on the grounds that dropout is one type of negative outcome.

Methods

Participants

A full description of the sample characteristics and treatment protocols, along with the main treatment outcome findings, has been reported elsewhere (DeRubeis et al., 2005; Hollon, DeRubeis et al., 2005). The sample of individuals randomized to cognitive therapy and antidepressant medications consisted of 180 depressed outpatients (measured using the Structured Clinical Interview for DSM-IV Diagnosis, SCID-I, First, Spitzer, Gibbon, & Williams, 2001) who scored at or above 20 on two successive interviews on the modified 17-item version of the Hamilton Rating Scale for Depression (HRSD: Hamilton, 1960) which were conducted one week apart. Exclusion criteria were: history of Bipolar I disorder, active substance abuse, psychosis, previous failed response to study medications, or the presence of another Axis I disorder that was judged to be primary. In addition, patients with antisocial, borderline, or schizotypal personality disorders were excluded. (Personality disorder diagnoses were made at intake using the Structured Clinical Interview for DSM-III-R Personality Disorders, SCID-II; Spitzer, Williams, Gibbon, & First, 1990.) Finally, patients were excluded if they were judged to be at such a high risk for suicide that immediate hospitalization was deemed necessary. Institutional review boards at each site approved the study protocol, and all patients provided written informed consent.

Treatments

Prior to entering the trial, patients were randomly assigned to receive cognitive therapy (N = 60) or antidepressant medication (N=120). An additional 60 patients were randomized to an 8-week pill-placebo condition. Patients and pharmacotherapists were blind to treatment condition during this 8-week period. Because data from the placebo arm of the study would not aid in the identification of prognostic or prescriptive variables for the two active treatments, these data were not included in the subsequent analyses. The primary medication used in the study was paroxetine, allowing for augmentation with lithium or desipramine if necessary. Acute treatment was provided for 16 weeks. There were twice the number of subjects in the medication condition because at the end of the acute treatment phase, medication responders

were randomized a second time for a follow-up study of relapse (see Hollon, DeRubeis et al., 2005). Data from the follow-up period will not be examined in this investigation.

Outcome Measures

The primary outcome measure was the 17-item version of the HRSD. Independent evaluators who were blind to treatment condition conducted all assessments, which were held once a week for the first four weeks, and once every two weeks thereafter.

Potential Predictors

All potential predictors of treatment response were measured at baseline, prior to randomization. To help control the number of predictors to be examined, and following a similar decision rule used by Morral and colleagues (Morral, Iguchi, Belding, & Lamb, 1997), a minimum distribution requirement was imposed for all dichotomous variables whereby the smaller group in the dichotomy had to comprise at least 10% of the sample in each treatment condition in order to be considered further. Each variable was assigned to one of the following five domains:

History of illness—This domain contained the following seven variables: *Number of prior episodes*, *Onset age*, *Chronic depression* (self-reported duration of the current episode ≥ 2 years), *Dysthymia* (assessed using the SCID-I interview), *Recurrent depression* (≥ 2 prior episodes of depression), *Atypical depression* and *Melancholic depression* (assessed using the SCID-I interview). Patients could be considered to have both atypical and melancholic features if they met the criteria for both).

Demographics and life circumstances—The following nine variables comprised this domain: *Age*, *Employment status*, *Gender*, *Marital status* (single vs. married or cohabitating), *Race* (because no other racial group met the minimum distribution requirement, the dichotomous Caucasian / non-Caucasian distinction was used), *Number of years of education*, *Income* (Gross annual income, measured in thousands of US dollars), *Total number of life events at baseline* (the total number of life events reported on the Psychiatric Epidemiology Research Interview Life Events scale, a 102-item self-report measure; Dohrenwend, Krasnoff, Askenasy, & Dohrenwend, 1978. This instrument is agnostic as to whether life events are positive, negative, or neutral), and *Intelligence* (Intelligence quotient, IQ, estimates were derived from the Shipley-Harford Living Scale; Shipley, 1940, a self-administered intelligence screen composed of a 20-item verbal intelligence scale and a 10-item analytic intelligence scale. Scores on these scales can be combined to estimate IQ scores on the Wechsler Adult Intelligence Scale – Revised; Zachary, 2001).

Family history of mental illness—To assess family history of mental illness, we used the Family History-Research Diagnostic Criteria, modified to yield DSM-IV diagnoses (Andreasen, Endicott, Spitzer, & Winokur, 1977). All patients were asked to provide information regarding the presence of up to two psychiatric diagnoses for first-degree relatives. To calculate the load of an individual patient's family history of mental illness, percentages were calculated by dividing the number of first degree relatives meeting a particular criterion by the total number of first degree relatives reported. We examined the following five variables reflecting the percentages of incidence among first-degree relatives: *Major Depressive Disorder*, *Any other mental disorder*, *Hospitalized for psychiatric reasons*, *Prescribed psychiatric medications*, and *Attempted suicide*.

Cognitive dysfunction—Five cognitive variables were included in the cognitive dysfunction domain: *Attributional style* (computed from the Attributional Styles Questionnaire, a 12-vignette self-report instrument composed of six positive and six negative

life events; Seligman, Abramson, Semmel, & von Baeyer, 1979. Participants rate the degree to which the causes of each event were internal, global, and stable. The composite composed of the difference between the internal, global, and stable ratings for positive and negative events was used for all analyses, DeRubeis et al., 1990.), *Perfectionism* and *Need for approval* (the two factors identified in previous work [Shahar, Blatt, Zuroff, & Pilkonis, 2003] to underlie the Dysfunctional Attitude Scale, a 40-item self report measure that indexes attitudes and beliefs hypothesized to underlie the thought processes of depressed individuals; Weissman & Beck, 1978), *Self-esteem* (assessed using the total score from the 10-item Rosenberg Self-Esteem Scale; Rosenberg, 1965), and *Hopelessness* (assessed using the total score from the 20-item, self report Hopelessness Scale; Beck, Weissman, Lester, & Trexler, 1974).

Baseline functioning—Twelve variables comprised the baseline functioning domain: *Beck-anxiety* (assessed using the self report Beck Anxiety Inventory; Beck, Epstein, Brown, & Steer, 1988), *Hamilton-anxiety* (assessed using the Hamilton Rating Scale for Anxiety, a clinician administered rating scale, Hamilton, 1959), *Anxiety sensitivity* (measured with the Anxiety Sensitivity Index, a 16-item self-report measure that assesses concerns about the consequences of anxious arousal, (Reiss, Peterson, Gursky, & McNally, 1986), *Global Assessment of Functioning* (GAF; American Psychiatric Association, 2000), *Axis I comorbidity* (any comorbid Axis-I disorder diagnosed using the SCID-I), *Positive affect* and *negative affect* (assessed using the Positive and Negative Affect Scale, PANAS, Watson, Clark, & Tellegen, 1988. The original PANAS was modified by adding 29 items designed to assess the circumplex model of emotion, Larsen & Diener, 1992; however, only the items composing the original PANAS were included). Finally, all five traits from the Five-Factor model of personality, *Agreeableness*, *Conscientiousness*, *Extraversion*, *Neuroticism*, and *Openness* were included (assessed using the 60 item NEO-Five-Factor Inventory, Costa & McCrae, 1992).

Data Analytic Strategy

Our primary statistical analyses investigated the association between possible prognostic and prescriptive predictors of symptom change across the 16-week study. Continuous data from the HRSD were analyzed using hierarchical linear models (HLM, also known as random regression models or growth curve models) that adjusted for the repeated measures with nested random effects. Using this approach, each subject's growth curve and HRSD score at the end of treatment is estimated from a collection of patient-specific parameters. These are treated as having been randomly sampled from a population of individuals, and an unstructured covariance structure was assumed in order to model the correlation between the patient-specific intercepts and slopes. All available data were used, rendering this data analytic strategy a full intent-to-treat analysis. Two baseline scores were obtained for each participant, allowing us to covary each patient's initial baseline depression severity score and to maintain a full intent-to-treat approach, even for individuals who dropped out before the first day of treatment. For all models, full maximum likelihood estimation procedures were used, and the degrees of freedom for hypothesis tests were estimated with the Kenward-Roger approximation (Kenward & Roger, 1997). All analyses were performed using SAS Version 9.1 PROC MIXED (SAS Institute Inc, Cary, NC).

To identify relevant predictors, we were guided by the approach advocated by Kraemer and colleagues (Kraemer et al., 2002). Within this framework, the interaction between treatment condition and the predictor of interest is examined. If the term representing this interaction is significantly related to outcome, the predictor is considered to be prescriptive, as it indicates differential effects of the treatments depending on the value of the variable in question. If the interaction term is not significant but the lower order term representing the effect of the variable is significant, then the predictor is considered to be prognostic. In such a case, outcome depends

on the level of this predictor independent of the treatment that was received. Using this general framework, we constructed the HLMs to assess simultaneously whether a variable was prognostic or prescriptive. To determine whether a variable was prognostic, we examined the effect of the predictor at the intercept (centered to represent estimated endpoint scores at 16 weeks) and on the linear slope estimates (represented in the model by the predictor-by-time interaction). Prognostic predictors were required to be associated with both of these outcomes (intercepts and slopes) at the $p < .05$ level. In order to determine whether the variable was prescriptive, we investigated the predictor-by-treatment interaction effect at the intercept as well for linear slope estimates. Prescriptive predictors were required to be associated with both of these outcomes (intercepts and slopes) at $p < .05$.

Given the potentially large number of statistical tests implied by the number of baseline variables to be examined, analyzing each separately would be expected to increase the likelihood that we would find significant relationships purely by chance. One typical solution to this dilemma is the imposition of a correction factor, such as the Bonferroni, which would raise the threshold required for declaring statistical significance. Rothman (1990), on the other hand, argues that the use of such correction factors is not only misguided given the fundamental task of empirical science, but might inadvertently lead to a greater number of errors of inference, albeit errors in the opposite direction. That is, such corrections might render significance tests to be so strict as to reject real relationships between variables that would have been worth exploring had they been examined separately. In an attempt to maintain a balance between these two competing concerns, we employed the following approach.

First, we calculated separate models for each of the aforementioned predictor domains. For each domain, a larger *prediction model* containing all of the potential predictors in that domain was compared to a smaller, nested model. This smaller, *simple model* contained the covariates implemented in the original (DeRubeis et al., 2005) manuscript (main effects of baseline HRSD, site, and treatment, as well as the site by treatment interaction). The fit of the prediction model was compared to that of the simple model by means of likelihood ratio tests between the models' deviance statistics (Singer & Willett, 2003). Only in cases where the prediction model proved statistically superior to the simple model at $\alpha \leq .05$ were the specific effects of the individual predictors explored. Because the deviance statistic requires that both the simple and prediction models contain all of the same individuals, missing data from any of the potential predictors could be problematic as it would result in the list-wise deletion of individuals with missing data from one model but not the other. Across most predictors, the rate of missing data was below 9%. The rates of missing data were slightly higher for some of the family history of mental illness predictors, the highest being 18% for the "family history of suicide attempts" variable. All missing values were imputed using multiple regression models containing site, gender, and age as predictors.

We employed a step-wise procedure within each of the five domains. Step 1 of the procedure for a given domain was the test of whether the model that included all variables from the domain was significant. In Step 2, we retained only those predictors associated in Step 1 with significance values of $p < .20$; in Step 3, we retained only those from Step 2 where $p < .10$. Finally, in Step 4, we retained only those predictors from Step 3 with a significance value of $p < .05$. As in previous work of this nature (Hybels, Blazer, & Steffens, 2005), once all predictor variables were identified, they were entered into a full model containing all significant predictors so that we could ascertain the effects of each variable while simultaneously controlling for each of the other significant predictors. For all models reported below, continuous variables were centered at the grand mean, and dichotomous variables were set to $-\frac{1}{2}$ and $\frac{1}{2}$ (Kraemer & Blasey, 2004).

Results

Omnibus tests of the five predictor domains

Of the five domains examined, only three of them fit the data better than the simple model, as indicated by significant likelihood ratio chi-squared tests: history of illness, $\chi^2(28) = 42.3$, $p = .041$, demographics and life circumstances, $\chi^2(36) = 62.0$, $p = .005$, and baseline functioning, $\chi^2(48) = 84.7$, $p < .001$ ³. The tests for family history, $\chi^2(20) = 20.1$, $p = 0.45$, and cognitive dysfunction, $\chi^2(20) = 27.1$, $p = 0.13$, were not significant.

Stepwise analyses of the three significant domains

History of illness—Of the seven variables from this domain, only one, the presence of chronic depression, emerged in Step 4 as a significant prognostic indicator of treatment outcome. It was significantly related to estimated HRSD scores at 16 weeks, $t(159) = 2.99$, $p = 0.003$, and to linear change in HRSD scores over time, $t(158) = 3.26$, $p = 0.001$. The results of the stepwise analysis are presented in Table 1. One of the criteria for the identification of prescriptive and prognostic predictors was that they prove significant both in the prediction of estimated Week 16 scores and linear slope estimates. However, for ease of presentation, we present only results from the prediction of Week 16 scores in this and subsequent tables. The estimates displayed in the tables can be interpreted as follows. For the lower-order effects, the unstandardized b estimates can be interpreted as representing the change in estimated Week 16 HRSD scores per unit of the predictor above the sample mean, controlling for all of the other predictors in the model. If the predictor is dichotomous, the estimate represents the difference between the two categories. For example, in Step 1 of the model, the b estimate for melancholia is -3.62 , indicating that individuals with melancholia can be expected to score 3.62 points lower on the HRSD at Week 16 than their non-melancholic counterparts. Similarly, the b estimates for the interaction terms represent the difference in the effect of the variables between the two treatments. For example, the b estimate of -5.52 for the melancholia-by-treatment interaction indicates that the magnitude of the advantage of melancholia over non-melancholia differs by an estimated 5.52 HRSD points between the two treatments. At 16 Weeks, melancholic individuals treated with antidepressant medications were estimated to score .86 HRSD points lower than non-melancholic individuals (with estimated scores of 6.81 and 7.67, respectively), whereas melancholic individuals receiving cognitive therapy were estimated to score 6.38 HRSD points lower than their non-melancholic counterparts (with estimated Week 16 scores of 2.69 and 9.07, respectively). Hence the difference in the magnitude of the melancholia effect between the two treatments is 5.52. As noted in Table 1, however, neither the main effect of melancholia nor its interaction with treatment reached the $p < .05$ criterion in Step 4, and hence, neither was retained in the final model.

Demographics and life circumstances—Table 2 displays the results of the analysis for the group of 10 demographic predictor variables. Two prognostic variables, age and intelligence, were associated with Week 16 scores ($t[159] = 2.59$, $p = 0.01$ for age; $t[160] = -1.99$, $p = 0.049$ for intelligence) as well as linear slope estimates ($t[159] = 2.98$, $p = .003$ for age; $t[161] = -2.14$, $p = 0.03$ for intelligence). In addition, three prescriptive variables emerged from this domain: marital status, employment status, and the number of life events reported at baseline. Each predicted a differential effect of treatment on Week 16 scores ($t[156] = -3.13$, $p = 0.002$ for marital status; $t[162] = -3.04$, $p = 0.003$ for employment status; and $t[158] = -2.17$, $p = .03$ for life events) and linear slope estimates ($t[156] = -2.86$, $p = .005$ for marital

³Twelve predictors belonged to the baseline functioning domain, and none was selected by the rules outlined above for inclusion in the final model. One variable, agreeableness, failed to meet the criteria, by the smallest of margins. It registered as a significant prognostic predictor of Week 16 scores, $t(163) = -2.70$, $p = 0.008$, however it failed to be significant predictor of linear slope estimates, $t(161) = -1.96$, $p = 0.052$. It was therefore not included in the final prediction model, in accordance with the criteria we detailed in the Methods section.

status; $t [163] = -3.04, p = 0.003$ for employment status; and $t [160] = -2.35, p = 0.02$ for life events).

Final model with all significant predictors

Once identified, the three prognostic and three prescriptive markers were entered simultaneously into a final model so that the effect of each could be ascertained while controlling for the effects of the others. As displayed in Table 3, each of the effects presented above remained significant when the effects of all of the other significant predictors were covaried. None of these pre-treatment variables differed significantly between the two treatments at intake; for the three continuous variables, all $t_s (178) \leq |0.39|$, all $p_s > 0.69$; for the three dichotomous variables, all $\chi^2_s (1, N = 180) \leq 1.62$, all $p_s > 0.20$.

Table 4 displays the estimated Week 16 scores for each of the prognostic predictors, averaged across the two treatments. As shown in the table, there is a positive relationship between age and end-of-treatment scores; each year of age above the mean age of the sample ($M = 40, SD = 12$) was associated with 0.10 additional residual symptoms on the HRSD at Week 16. Conversely, there is a negative relationship between end-of-treatment scores and intelligence. Each IQ point greater than the mean of the sample ($M = 109, SD = 11$) was associated with 0.10 fewer HRSD points at 16 Weeks.

For the three prescriptive predictors of treatment outcome, separate estimates of Week 16 HRSD scores were calculated for each treatment group. Figure 1 displays estimated end-of-treatment scores by marital status in each of the two treatment conditions. Although no difference is evident between the two treatments for unmarried participants, $t(160) = -0.05, p = 0.96$, Cohen's $d = -0.01 \pm 0.38$ (95% CI), married or cohabiting participants evidenced lower end of treatment scores following cognitive therapy, relative to antidepressant medication, $t(160) = 3.70, p < 0.001$, Cohen's $d = 1.04 \pm 0.58$ (95% CI). (Effect size estimates were calculated from least-squares means estimates of week 16 scores.)

Figure 2 displays the differential effect of employment status on outcome in the two treatments. For participants who were employed, there was no difference between the two treatments $t (155) = -0.67, p = 0.51$, Cohen's $d = -0.12 \pm 0.35$ (95% CI), however, for unemployed participants, cognitive therapy was associated with superior outcomes relative to medication, $t(163) = 3.24, p = 0.002$, Cohen's $d = 1.19 \pm 0.78$ (95% CI).

Finally, Figure 3 displays end-of-treatment HRSD score estimates for each treatment modality for three levels of baseline life events: the sample mean, one standard deviation above the mean, and one standard deviation below the mean ($M = 7.0, SD = 5.3$). As displayed in the figure, there was no difference between the two treatments in estimated Week 16 scores at one standard deviation below the mean number of life events, $t(155) = 0.80, p = 0.43$, Cohen's $d = 0.13 \pm 0.31$ (95% CI). As the number of reported life events increased, participants in the cognitive therapy condition were estimated to display fewer symptoms at the end of treatment, whereas individuals who had received medications were predicted to show a greater number of symptoms at the end of treatment. The estimated Week 16 scores differed between the two treatments both at the mean number of baseline life events, $t(162) = 2.65, p = 0.009$, Cohen's $d = 0.42 \pm 0.31$ (95% CI), and at one standard deviation above the mean, $t(164) = 3.11, p = 0.002$, Cohen's $d = 0.49 \pm 0.31$ (95% CI).¹

The final model containing the three prognostic and three prescriptive variables accounted for approximately 33% of the between-subjects variance in both estimated Week 16 scores and linear slope estimates. When the two prescriptive factors identified in prior publications from the same dataset – personality disorder status and prior medication treatment – were entered into the final model, the percent of between-subjects variance explained by the model increased

to 41% for the Week 16 scores and 42% for the linear slope estimates. The addition of these two variables did not substantially alter the results reported above.²

Predicting Dropout

In a secondary set of analyses we attempted to predict rates of attrition. We used the same analytical framework outlined above, adapted for use with logistic regression models based on the Likelihood Ratio Chi-square statistic (Hosmer & Lemeshow, 1989). Because of the small number of dropouts relative to the number of potential predictors, we could not investigate prescriptive predictors of dropout. Had we done so, we would have been forced by our analytic strategy to test models with more predictors than there were dropout events to predict.

The overall rates of attrition were similar in the two treatments (16% for medications and 17% for cognitive therapy; $\chi^2(1, N = 180) = 0.02$, $p = 0.89$). The simple, nested model used to test each of the five domains was composed of only two variables, site and treatment condition. Using the procedures outlined above, only two significant predictors of dropout emerged: history of chronic depression, $\chi^2(1, N = 180) = 7.19$, $p = 0.007$, and the total number of life events at baseline, $\chi^2(1, N = 180) = 4.97$, $p = 0.03$. Specifically, dropout rates were lower for individuals with chronic depression. Of those with chronic depression, only 8% (95% CI: 2–13%) of participants dropped out, compared to 25% (95% CI: 16–34%) of those with non-chronic forms of the illness. This corresponds to an odds ratio of 4.0 (95% CI: 1.6–10.0%), indicating that individuals with non-chronic depression were approximately four times more likely to dropout compared to those with chronic depression. Regarding life events at baseline, Figure 4 depicts a pattern whereby as the number of pre-treatment life events increased, so too did the probability that an individual would drop out of treatment. Because both chronic depression and baseline life events also emerged as predictors of outcome in the analyses reported above, in order to test whether their emergence as predictors of outcome could be accounted for by the fact that they predicted attrition, we conducted completers-only analyses of outcome for each of these two variables. The same pattern of results reported from the intent-to-treat analyses were obtained for the completers-only analyses as well.

¹In each of Figure 1, Figure 2 and Figure 3 there appears to be an overall difference between medications and cognitive therapy, with the latter yielding lower symptom scores overall. Indeed, in the final model that included all significant predictors, the main effect of treatment was significant for estimated Week 16 scores, $t(162) = -2.65$, $p = 0.009$ as well as linear slope estimates, $t(162) = -2.34$, $p = 0.02$. These effects, which were not evident under the simpler model that was the basis for the original DeRubeis et al. (2005) report, must be interpreted in the proper context. They appear to arise from the fact that the proportion of married participants in the sample departed substantially from expected population rates of 50% (it was 34%), as did the proportion of unemployed participants (17%). There were no differences between the two treatments in the distributions of these variables; $\chi^2[1, N = 180] = 0.61$, $p = 0.44$ for marital status; $\chi^2[1, N = 180] = 1.62$, $p = 0.20$ for employment status. In both cases, cognitive therapy was more effective for members of the minority group. In all models presented in this paper, unweighted effects coding ($-\frac{1}{2}$ and $\frac{1}{2}$) was used for dichotomous variables, which results in parameter estimates expected to generalize to a population wherein each constituent of the dichotomy is expected to occur in 50% of individuals. When the marital status and employment status variables were coded with a weighted effects coding system that reflects the proportions observed in our sample, the main effect of treatment was no longer significant, $t(158) = 0.47$, $p = 0.64$ for Week 16 scores; $t(159) = 0.83$, $p = 0.41$ for linear slope estimates), paralleling the results of the analyses reported in DeRubeis et al. The choice of coding system does not affect t -statistics or associated p -values for any of the prognostic or prescriptive terms on which we focus in this study. Because weighted effects coding makes the strong claim that the proportions we observed in our sample reflect those in the larger population, we chose to present results using the unweighted effects coding system. This same assumption is the default in most statistical ANOVA software packages.

²When all of the significant predictors identified in the present study, as well as the two variables identified in previous work, were added to the same model, the direction of all results remained unchanged, as did the significance level of the majority of the statistical tests. All variables identified in previous manuscripts remained significant in the prediction of Week 16 scores, as did the majority of those identified in the present effort. Accounting for the effects of the previously identified variables, the main effect of Age and IQ on estimated Week 16 scores dropped to the level of non-significant trends ($t(160) = 1.73$, $p = 0.08$ for the effect of age, and $t(160) = -1.70$, $p = 0.09$, for the effect IQ). Such findings are perhaps not surprising given that the inclusion of each new variable into the model is associated with a reduction in power for the hypothesis tests of the other variables.

Discussion

We sought to identify predictors of response for two well-validated treatments of depression. In particular, our aim was to identify prognostic variables as well as variables that, if replicated, could be used to make prescriptive recommendations. One of the variables that predicted poorer response across the two treatments was age, which has been previously identified in a mega-analysis conducted by Thase and colleagues (1997) as a predictor of slower recovery to treatment for depression. In at least five other investigations, however, age has not been found to predict treatment outcomes (Hirschfeld et al., 1998; Petersen et al., 2002; Sotsky et al., 1991; Szadoczký et al., 2004; Tuma, 1996). Similarly, investigators have not found intelligence to be a prognostic factor (Dunkin et al., 2000; Haaga, DeRubeis, Stewart, & Beck, 1991), although it was a significant prognostic indicator in the present study.

As has been found in previous investigations (Blom et al., 2007; Joyce et al., 2002; Sotsky et al., 1991; Thase et al., 1994), chronicity emerged as a negative prognostic indicator. Although chronically depressed patients experienced less symptom relief, on average, than did non-chronic patients in the present study, they were more likely to complete the 16-week treatment protocol. One possible explanation for this apparent discrepancy is that chronically depressed individuals were more highly motivated to remain in treatment with the hope of achieving some degree of relief from the long-standing illness from which they had been suffering. The fact that their condition was chronic is, on its own, evidence that the symptoms were not likely to be remediated easily or rapidly.

Investigators constructing future studies of this sort, in which medications and cognitive therapy are compared directly, might consider stratifying random assignment of individuals to treatment conditions based on these three prognostic variables so as not to inadvertently bias the results in favor of one treatment over the other (Nierenberg, 2003). Should efforts to replicate the three prognostic variables prove successful, additional research work might attempt to identify other therapeutic strategies or modalities that might be more effective at treating individuals with these characteristics.

In addition, we identified three baseline variables that predicted differential response to the two treatments. Specifically, cognitive therapy was more efficacious than antidepressant medications for those experiencing a greater number of events in their lives, as well as for those who were married or cohabiting, and for those who were unemployed. Findings from previous predictive analyses regarding baseline life events have been mixed, with some finding that higher levels of stressful life events predict poorer response to antidepressant treatment (Mazure et al., 2000; Monroe, Kupfer, & Frank, 1992); whereas others have found either the reverse pattern (Monroe, Bellack, Hersen, & Himmelhoch, 1983) or no relationship (Szadoczký et al., 2004). In a study examining differences in response between cognitive therapy and antidepressant medications in the TDCRP sample, Sotsky and colleagues (1991) found no evidence that marital status predicted differential response. However, in a follow-up analysis examining only the two psychotherapy conditions in the TDCRP, cognitive therapy and interpersonal therapy, Barber and Muenz (1996) reported that cognitive therapy was more effective at reducing depressive symptoms for married or cohabiting individuals, whereas nonmarried individuals responded more favorably following interpersonal therapy.

Should the prescriptive effects identified herein be replicated, the mechanisms by which these moderators exert their effects should be examined in future research. Kazdin and Nock (2003) note that moderation, referred to in this paper as the identification of prescriptive variables, implies that there are different mechanisms involved in the respective treatment modalities. It is possible, for example, that individuals who are married or cohabiting, unemployed, or experiencing a greater number of life events might be particularly suited to

cognitive therapy in that they might present to treatment with a number of easily identifiable stressors to which the tools they learn in therapy could be readily applied. The mechanism(s) of action for antidepressant medications, on the other hand, might not remediate the cognitive/neural/behavioral systems that are impacted by these stressors.

Limitations

There are a number of limitations inherent in the current study. First, our analytic strategy was one that attempted to traverse a middle path between type-I and type-II error rates. We therefore did not minimize either type of error to the extent that might occur in a strict hypothesis testing study or an exploratory investigation, respectively. Had we instead conducted separate hypothesis tests for each of the potential predictors for both Week 16 scores and linear slope estimates, and examined prognostic and prescriptive effects separately (either in a single report or in several separate reports), we would have conducted approximately 150 hypothesis tests, from which we could have expected to find several statistically significant results by chance. On the other hand, had we set the overall type-I error rate to .05, we would have certainly missed potentially meaningful relationships, and the respective pre-treatment variables would then be unlikely to be included in future prediction studies or in attempts to aggregate findings across investigations, as in a meta-analysis.

Second, given the nature of the study from which these data originate, the results reported herein can only be expected, strictly speaking, to generalize to outpatients diagnosed with moderate to severe depression treated for 16 weeks with the therapeutic modalities employed in this study. It is not clear from this study whether different antidepressant medications or a longer duration of treatment might have led to different relations between pre-treatment variables and response. The robustness of the effects of the six predictors identified in this study might be examined in future work in which different methodological features are present. Additionally, the efficacy of treatment modalities other than cognitive therapy and the antidepressants used in this study might be investigated, particularly for those patients who fared poorly in the two treatments used in the current study.

A third limitation of this study is the intentional omission of interaction effects between the identified predictors. It is possible, for example, that the relationship between marital status and outcome might itself be moderated by other variables, such as age. Given the number of significance tests that were examined in this study, we believed that exploring such interactions would have unnecessarily increased the likelihood of capitalizing on chance. Provided that future targeted studies successfully replicate the results of this investigation, it might be more appropriate in such studies for more complicated relationships to be explored.

Finally, the selection of the candidate variables used in this study was made on the basis of prior reports in the literature. Historically some variables, such as the cognitive variables, have been chosen for theoretical reasons. However, others likely have been chosen because they were readily available to researchers. The sparseness of theory in this research area to date has likely contributed to the difficulty that researchers have experienced in replicating findings across settings. The statistical approach used in this manuscript, however, may render the results reported herein more likely to replicate than those reported in previous work in this area. If these results do replicate, they might be used as seeds to help to develop theories about the nature of depression and its treatment, from which future hypotheses could be formulated.

Clinical implications

The pattern of findings from this report, if replicated, could suggest clinically meaningful recommendations for treatment providers. The three prognostic variables, age, intelligence, and chronicity, might be used to identify those patients for whom a longer duration or higher

dose of treatment is likely to be needed. Alternatively, it is possible that these individuals might benefit more than their counterparts from a combination of cognitive therapy and medications, or from a treatment modality not included in this investigation. For depressed individuals who are married or cohabiting, unemployed, or experiencing a large number of events in their lives, the results of this study suggest that cognitive therapy might be considered to be the treatment of choice, quite aside from its long-term benefits (Hollon, DeRubeis et al., 2005) as it may lead to a substantially greater reduction in symptoms after 16 weeks of treatment than will the medications used in this study.

Acknowledgments

We would like to thank our colleagues for making this research possible. Paula R. Young and Margaret L. Lovett served as the two study coordinators. John P. O'Reardon, Ronald M. Salomon and the late Martin Szuba served as study pharmacotherapists (along with J.D.A and R.C.S.). Cory P. Newman, Karl N. Jannasch, Frances Shusman and Sandra Seidel served as the cognitive therapists (along with R.J.D. and S.D.H.). Jan Fawcett provided consultation with regard to the implementation of clinical management pharmacotherapy. Aaron T. Beck, Judith Beck, Christine Johnson and Leslie Sokol provided consultation with respect to the implementation of cognitive therapy. Madeline M. Gladis and Kirsten L. Haman oversaw the training of the clinical interviewers, and David Appelbaum, Laurel L. Brown, Richard C. Carson, Barrie Franklin, Nana A. Landenberger, Jessica Londa-Jacobs, Julie L. Pickholtz, Pamela Fawcett-Pressman, Sabine Schmid, Ellen D. Stoddard, Michael Suminski and Dorothy Tucker served as the clinical interviewers. Joyce L. Bell, Brent B. Freeman, Cara C. Grugan, Nathaniel R. Herr, Mary B. Hooper, Miriam Hundert, Veni Linos and Tynya Patton provided research support..

This research was supported by grants MH50129 (R10), MH55875 (R10), MH01697 (K02), and MH01741 (K24) from the National Institute of Mental Health, Bethesda, Maryland, USA. GlaxoSmithKline (Brentford, U.K.) provided medications and pill placebos for the trial.

References

- Agosti V, Ocepek-Welikson K. The efficacy of imipramine and psychotherapy in early-onset chronic depression: A reanalysis of the National Institute of Mental Health Treatment of Depression Collaborative Research Program. *Journal of Affective Disorders* 1997;43:181–186. [PubMed: 9186788]
- Aguinis H, Stone-Romero EF. Methodological artifacts in moderated multiple regression and their effects on statistical power. *Journal of Applied Psychology* 1997;82:192–206.
- Aiken, LS.; West, SG. *Multiple Regression: Testing and Interpreting Interactions*. Newbury Park CA: Sage; 1991.
- Andreasen NC, Endicott J, Spitzer RL, Winokur G. The family history method using diagnostic criteria: Reliability and validity. *Archives of General Psychiatry* 1977;34:1229–1235. [PubMed: 911222]
- American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*. Vol. 4th ed., text revision ed.. Washington, DC: American Psychiatric Press; 2000.
- Barber JP, Muenz LR. The role of avoidance and obsessiveness in matching patients to cognitive and interpersonal psychotherapy: Empirical findings from the Treatment for Depression Collaborative Research Program. *Journal of Consulting and Clinical Psychology* 1996;64:951–958. [PubMed: 8916624]
- Beck AT, Epstein N, Brown G, Steer RA. An inventory for measuring clinical anxiety: Psychometric properties. *Journal of Consulting and Clinical Psychology* 1988;56:893–897. [PubMed: 3204199]
- Beck AT, Weissman A, Lester D, Trexler L. The measurement of pessimism: The Hopelessness Scale. *Journal of Consulting and Clinical Psychology* 1974;42:861–865. [PubMed: 4436473]
- Blackburn IM, Moore RG. Controlled acute and follow-up trial of cognitive therapy and pharmacotherapy in out-patients with recurrent depression. *British Journal of Psychiatry* 1997;171:328–334. [PubMed: 9373420]
- Blatt SJ, Quinlan DM, Pilkonis PA, Shea MT. Impact of perfectionism and need for approval on the brief treatment of depression: The National Institute of Mental Health Treatment of Depression Collaborative Research Program revisited. *Journal of Consulting and Clinical Psychology* 1995;63:125–132. [PubMed: 7896977]

- Blatt SJ, Zuroff DC, Bondi CM, Sanislow CA, Pilkonis PA. When and how perfectionism impedes the brief treatment of depression: Further analyses of the National Institute of Mental Health Treatment of Depression Collaborative Research Program. *Journal of Consulting and Clinical Psychology* 1998;66:423–428. [PubMed: 9583345]
- Blom MJB, Spinhoven P, Hoffman T, Jonker K, Hoencamp E, Haffmans PMJ, et al. Severity and duration of depression, not personality factors, predict short term outcome in the treatment of major depression. *Journal of Affective Disorders* 2007;104:119–126. [PubMed: 17467059]
- Costa, PT.; McCrae, RR. Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) Professional Manual. Odessa, FL: Psychological Assessment Resources; 1992.
- DeRubeis RJ, Evans MD, Hollon SD, Garvey MJ, Grove WM, Tuason VB. How does cognitive therapy work? Cognitive change and symptom change in cognitive therapy and pharmacotherapy for depression. *Journal of Consulting and Clinical Psychology* 1990;58:862–869. [PubMed: 2292637]
- DeRubeis RJ, Gelfand LA, Tang TZ, Simons AD. Medications versus cognitive behavior therapy for severely depressed outpatients: Mega-analysis of four randomized comparisons. *American Journal of Psychiatry* 1999;156:1007–1013. [PubMed: 10401443]
- 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. *Archives of General Psychiatry* 2005;62:409–416. [PubMed: 15809408]
- Dohrenwend BS, Krasnoff L, Askenasy AR, Dohrenwend BP. Exemplification of a method for scaling life events: The PERI Life Events Scale. *Journal of Health and Social Behavior* 1978;19:205–229. [PubMed: 681735]
- Dunkin JJ, Leuchter AF, Cook IA, Kasl-Godley JE, Abrams M, Rosenberg-Thompson S. Executive dysfunction predicts nonresponse to fluoxetine in major depression. *Journal of Affective Disorders* 2000;60:13–23. [PubMed: 10940443]
- Elkin I, Gibbons RD, Shea MT, Sotsky SM, Watkins JT, Pilkonis PA, et al. Initial severity and differential treatment outcome in the National Institute of Mental Health Treatment of Depression Collaborative Research Program. *Journal of Consulting and Clinical Psychology* 1995;63:841–847. [PubMed: 7593878]
- Esposito K, Goodnick P. Predictors of response in depression. *Psychiatric Clinics of North America*. Special Issue: Drug therapy: Predictors of response 2003;26:353–365.
- Fava M, Uebelacker LA, Alpert JE, Nierenberg AA, Pava JA, Rosenbaum JF. Major depressive subtypes and treatment response. *Biological Psychiatry* 1997;42:568–576. [PubMed: 9376453]
- First, M.; Spitzer, R.; Gibbon, M.; Williams, J. Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Patient Edition With Psychotic Screen (SCID-I/P W/ PSY SCREEN). New York, NY: Biometrics Research, New York State Psychiatric Institute; 2001.
- Fournier JC, DeRubeis RJ, Shelton RC, Gallop R, Amsterdam JD, Hollon SD. Antidepressant medications v. cognitive therapy in people with depression with or without personality disorder. *British Journal of Psychiatry* 2008;192:124–129. [PubMed: 18245030]
- Goodwin FK. Predictors of antidepressant response. *Bulletin of the Menninger Clinic* 1993;57:146–160. [PubMed: 8508153]
- Haaga DA, DeRubeis RJ, Stewart BL, Beck AT. Relationship of intelligence with cognitive therapy outcome. *Behaviour Research and Therapy* 1991;29:277–281. [PubMed: 1883308]
- Hamilton M. The assessment of anxiety status by rating. *British Journal of Medical Psychology* 1959;32:50–55. [PubMed: 13638508]
- Hamilton MA. A rating scale for depression. *Journal of Neurology, Neurosurgery, and Psychiatry* 1960;23:56–62.
- Heiligenstein JH, Tollefson GD, Faries DE. Response patterns of depressed outpatients with and without melancholia: A double-blind, placebo-controlled trial of fluoxetine versus placebo. *Journal of Affective Disorders* 1994;30:163–173. [PubMed: 8006243]
- Hirschfeld RMA, Russell JM, Delgado PL, Fawcett J, Friedman RA, Harrison WM, et al. Predictors of response to acute treatment of chronic and double depression with sertraline or imipramine. *Journal of Clinical Psychiatry* 1998;59:669–675. [PubMed: 9921701]

- Hollon, SD.; Beck, AT. Predicting outcome vs. differential response: matching clients to treatment. Paper presented at the National Institute on Drug Abuse Technical Review Meeting on Matching Clients to Treatment: A Critical Review; Rockville, MD. 1986.
- Hollon SD, DeRubeis RJ, Shelton RC, Amsterdam JD, Salomon RM, O'Reardon JP, et al. Prevention of relapse following cognitive therapy vs medications in moderate to severe depression. *Archives of General Psychiatry* 2005;62:417–422. [PubMed: 15809409]
- 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? *Journal of Clinical Psychiatry* 2005;66:455–468. [PubMed: 15816788]
- Hosmer, D.; Lemeshow, S. *Applied Logistic Regression*. New York: Wiley; 1989.
- Hybels CF, Blazer DG, Steffens DC. Predictors of partial remission in older patients treated for major depression: The role of comorbid dysthymia. *American Journal of Geriatric Psychiatry* 2005;13:713–721. [PubMed: 16085788]
- Jacobson NS, Hollon SD. Cognitive-behavior therapy versus pharmacotherapy: Now that the jury's returned its verdict, it's time to present the rest of the evidence. *Journal of Consulting and Clinical Psychology* 1996;64:74–80. [PubMed: 8907086]
- Jarrett RB, Schaffer M, McIntire D, Witt-Browder A, Kraft D, Risser RC. Treatment of atypical depression with cognitive therapy or phenelzine: A double-blind, placebo-controlled trial. *Archives of General Psychiatry* 1999;56:431–437. [PubMed: 10232298]
- Joyce PR, Mulder RT, Luty SE, Sullivan PF, McKenzie JM, Abbott RM, et al. Patterns and predictors of remission, response and recovery in major depression treated with fluoxetine or nortriptyline. *Australian and New Zealand Journal of Psychiatry* 2002;36:384–391. [PubMed: 12060188]
- Kazdin AE, Nock MK. Delineating mechanisms of change in child and adolescent therapy: Methodological issues and research recommendations. *Journal of Child Psychology and Psychiatry* 2003;44:1116–1129. [PubMed: 14626454]
- Kenward MG, Roger JH. Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics* 1997;53:983–997. [PubMed: 9333350]
- Kraemer HC, Blasey CM. Centering in regression analyses: A strategy to prevent errors in statistical inference. *International Journal of Methods in Psychiatric Research* 2004;13:141–151. [PubMed: 15297898]
- Kraemer HC, Wilson GT, Fairburn CG, Agras WS. Mediators and moderators of treatment effects in randomized clinical trials. *Archives of General Psychiatry* 2002;59:877–884. [PubMed: 12365874]
- Larsen, RJ.; Diener, E. Promises and problems with the circumplex model of emotion. In: Clark, MS., editor. *Review of personality and social psychology: Emotion*. Vol. 13. Newbury Park, CA: Sage; 1992. p. 25–59.
- Leykin Y, Amsterdam JD, DeRubeis RJ, Gallop R, Shelton RC, Hollon SD. Progressive resistance to a selective serotonin reuptake inhibitor but not to cognitive therapy in the treatment of major depression. *Journal of Consulting and Clinical Psychology* 2007;75:267–276. [PubMed: 17469884]
- Mazure CM, Bruce ML, Maciejewski PK, Jacobs SC. Adverse life events and cognitive-personality characteristics in the prediction of major depression and antidepressant response. *American Journal of Psychiatry* 2000;157:896–903. [PubMed: 10831468]
- Monroe SM, Bellack AS, Hersen M, Himmelhoch JM. Life events, symptom course, and treatment outcome in unipolar depressed women. *Journal of Consulting and Clinical Psychology* 1983;51:604–615. [PubMed: 6619369]
- Monroe SM, Kupfer DJ, Frank EF. Life stress and treatment course of recurrent depression: I. Response during index episode. *Journal of Consulting and Clinical Psychology* 1992;60:718–724. [PubMed: 1401387]
- Morrall AR, Iguchi MY, Belding MA, Lamb RJ. Natural classes of treatment response. *Journal of Consulting and Clinical Psychology* 1997;65:673–685. [PubMed: 9256569]
- Nelson JC, Mazure CM, Jatlow PI. Does melancholia predict response in major depression? *Journal of Affective Disorders* 1990;18:157–165. [PubMed: 2139060]
- Nierenberg AA. Predictors of response to antidepressants: General principles and clinical implications. *Psychiatric Clinics of North America*. Special Issue: Drug therapy: Predictors of response 2003;26:345–352.

- Petersen T, Papakostas GI, Bottonari K, Iacoviello B, Alpert JE, Fava M, et al. NEO-FFI factor scores as predictors of clinical response to fluoxetine in depressed outpatients. *Psychiatry Research* 2002;109:9–16. [PubMed: 11850046]
- Raudenbush, SW.; Bryk, AS. *Hierarchical Linear Models: Applications and Data Analysis Methods*. Vol. 2nd. ed.. London: Sage Publication; 2002.
- Reiss S, Peterson RA, Gursky M, McNally RJ. Anxiety, sensitivity, anxiety frequency, and the prediction of fearfulness. *Behaviour Research and Therapy* 1986;24:1–8. [PubMed: 3947307]
- Rosenberg, M. *Society and the adolescent child*. Princeton, NJ: Princeton University Press; 1965.
- Rothman KJ. No adjustments are needed for multiple comparisons. *Epidemiology* 1990;1:43–46. [PubMed: 2081237]
- Rude SS, Rehm LP. Response to treatments for depression: The role of initial status on targeted cognitive and behavioral skills. *Clinical Psychology Review* 1991;11:493–514.
- Seligman ME, Abramson LY, Semmel A, von Baeyer C. Depressive attributional style. *Journal of Abnormal Psychology* 1979;88:242–247. [PubMed: 500951]
- Shahar G, Blatt SJ, Zuroff DC, Pilkonis PA. Role of perfectionism and personality disorder features in response to brief treatment for depression. *Journal of Consulting and Clinical Psychology* 2003;71:629–633. [PubMed: 12795586]
- Shipley WC. A self-administering scale for measuring intellectual impairment and deterioration. *Journal of Psychology* 1940;9:371–377.
- Singer, JD.; Willett, JB. *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence*. New York: Oxford University Press; 2003.
- Sotsky SM, Glass DR, Shea MT, Pilkonis PA, Collins JF, Elkin I, et al. Patient predictors of response to psychotherapy and pharmacotherapy: Findings in the NIMH Treatment of Depression Collaborative Research Program. *American Journal of Psychiatry* 1991;148:997–1008. [PubMed: 1853989]
- Spitzer, R.; Williams, J.; Gibbon, M.; First, M. *Structured Clinical Interview for DSMII-R Personality Disorders (SCID-II, Version 1.0)*. Washington, DC: American Psychiatric Press; 1990.
- Szadoczky E, Rozsa S, Zambori J, Furedi J. Predictors for 2-year outcome of major depressive episode. *Journal of Affective Disorders* 2004;83:49–57. [PubMed: 15546645]
- Thase ME, Greenhouse JB, Frank E, Reynolds CF III, Pilkonis PA, Hurley K, et al. Treatment of major depression with psychotherapy or psychotherapy-pharmacotherapy combinations. *Archives of General Psychiatry* 1997;54:1009–1015. [PubMed: 9366657]
- Thase ME, Reynolds CF, Frank E, Simons AD, Garamoni G, McGeary J. Response to cognitive-behavioral therapy in chronic depression. *Journal of Psychotherapy Practice & Research* 1994;3:204–214.
- Trivedi M, Morris DW, Grannemann BD, Mahadi S. Symptom clusters as predictors of late response to antidepressant treatment. *Journal of Clinical Psychiatry* 2005;66:1064–1070. [PubMed: 16086624]
- Trivedi MH, Rush AJ, Wisniewski SR, Nierenberg AA, Warden D, Ritz L, et al. Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: Implications for clinical practice. *American Journal of Psychiatry* 2006;163:28–40. [PubMed: 16390886]
- Tuma TA. Effect of age on the outcome of hospital treated depression. *British Journal of Psychiatry* 1996;168:76–81. [PubMed: 8770433]
- Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology* 1988;54:1063–1070. [PubMed: 3397865]
- Weissman, AN.; Beck, AT. Development and validation of the dysfunctional attitude scale: A preliminary investigation. Paper presented at the meeting of the American Educational Research Association; Toronto, Canada. 1978.
- Zachary, RA. *Shipley Institute of Living Scale: Revised Manual*. Los Angeles, CA: Western Psychological Services; 2001.

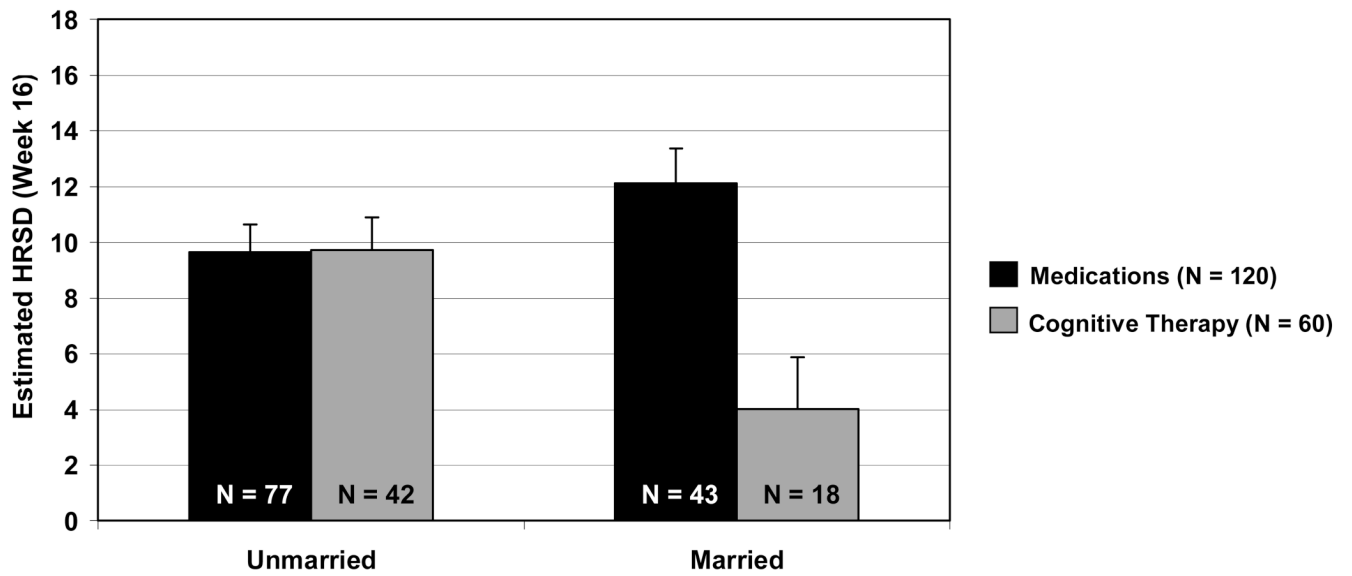


Figure 1. Estimated Week 16 HRSD (Hamilton Rating Scale for Depression) scores for married and unmarried participants in each of the two treatments. Error bars represent +/-1 standard error.

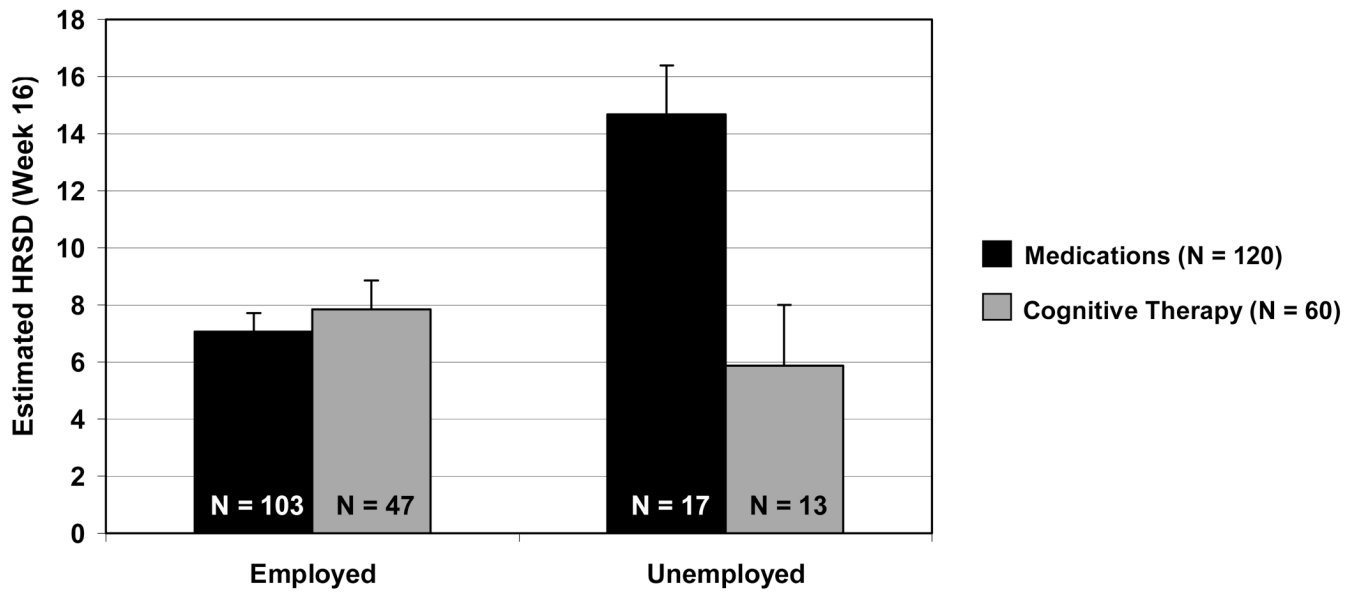


Figure 2. Estimated Week 16 HRSD (Hamilton Rating Scale for Depression) scores for employed and unemployed participants in each of the two treatments. Error bars represent +/-1 standard error.

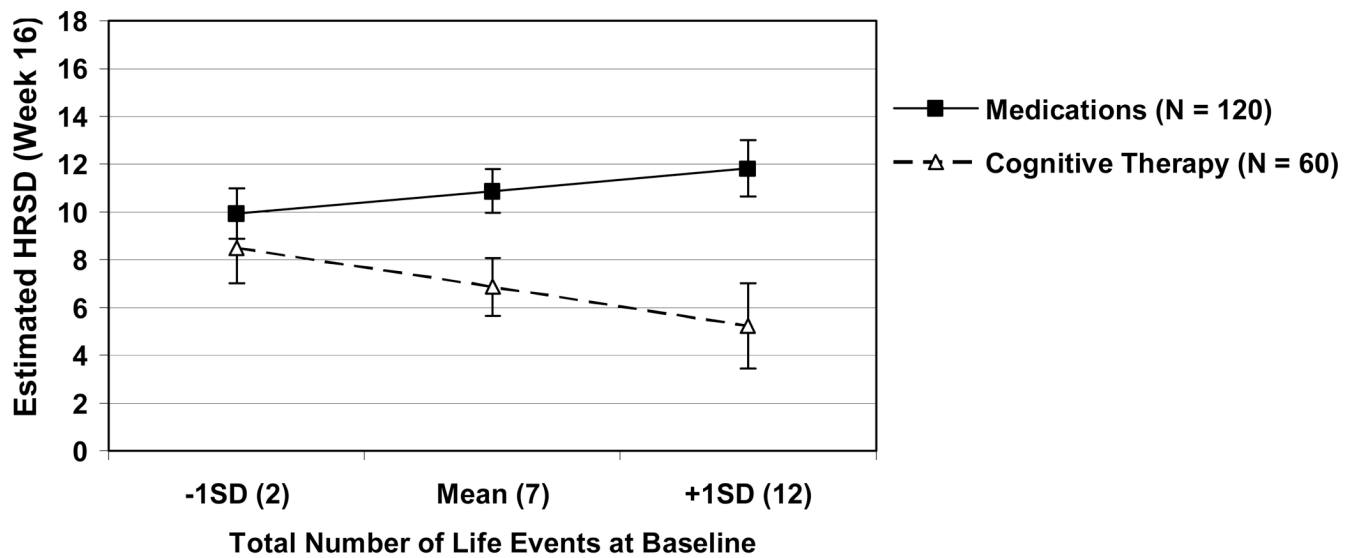


Figure 3.

Estimated Week 16 HRSD (Hamilton Rating Scale for Depression) scores for each treatment at each of three values of the total number of life events reported at baseline: the mean number of life events reported in the sample ($M=7$, $SD=5$), one standard deviation below the mean (2), and one standard deviation above the mean (12). Error bars represent +/-1 standard error.

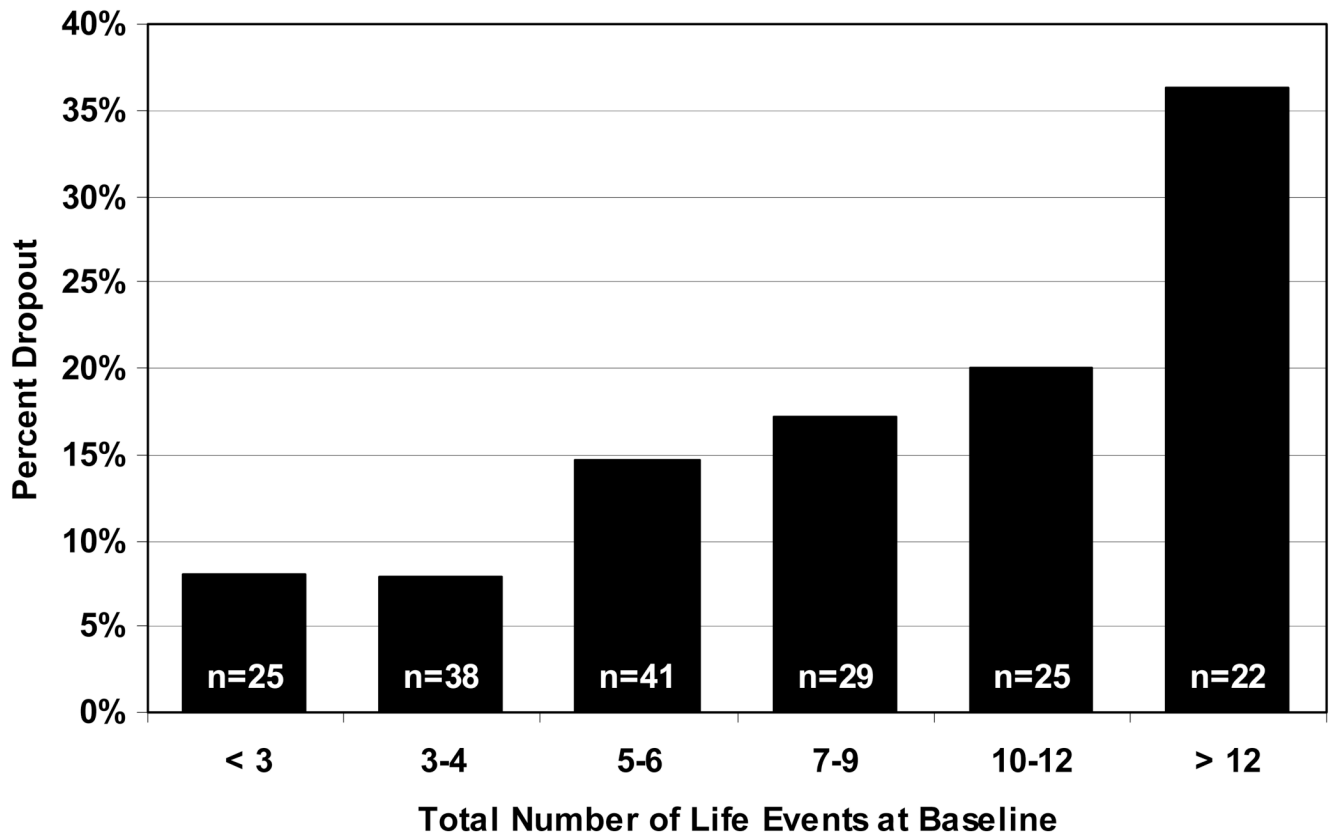


Figure 4. The percentage of participants who dropped out of treatment for each of six levels of total life events reported at baseline ($M=7$, $SD=5$).

Table 1

History of Illness

Step 1	Predictor	Mean (SD)	b^a	
		%		
	Recurrent MDD	73%	-1.45 (1.54)	
	Chronic MDD	51%	3.14 (1.19)	**
	Dysthymia	27%	-1.78 (1.41)	
	# Prior Episodes	2.48 (2.77)	0.16 (0.31)	
	Onset age	22.22 (12.22)	-0.03 (0.05)	
	Atypical	23%	-1.30 (1.37)	
	Melancholia	31%	-3.62 (1.52)	*
	Recurrent MDD × Treatment		-0.26 (3.09)	
	Chronic MDD × Treatment		-1.91 (2.38)	
	Dysthymia × Treatment		-0.24 (2.81)	
	# Prior Episodes × Treatment		-0.36 (0.61)	
	Onset age × Treatment		-0.12 (0.11)	
	Atypical × Treatment		-0.50 (2.75)	
	Melancholia × Treatment		-5.52 (3.05)	†
Step 2 (retain effects at $p < .20$)				
	Chronic MDD		3.26 (1.08)	**
	Melancholia		-3.17 (1.51)	*
	Melancholia × Treatment		-5.46 (3.02)	†
Step 3 (retain effects at $p < .10$)				
	Chronic MDD		3.26 (1.08)	**
	Melancholia		-3.17 (1.51)	*
	Melancholia × Treatment		-5.46 (3.02)	†
Step 4 (retain effects at $p < .05$)				
<u>Non significant</u>				
	Melancholia		-1.60 (1.24)	
<u>Significant</u>				
	Chronic MDD		3.26 (1.09)	**

Note. MDD = Major Depressive Disorder

^aFor this and subsequent tables, values represent unstandardized beta coefficients predicting HRSD scores at the end of acute treatment (Week 16 Scores). For lower-order terms, values represent effects averaged across the two treatment groups. Negative values represent more desirable outcomes (i.e., lower estimated week 16 scores). For interaction terms, values represent the difference between the two treatments in the magnitude of the effect. Parenthetical entries represent standard error estimates. The significant predictor in the final step will be added to the final model containing all significant predictors from each domain.

\ddagger p < .20

\dagger p < .10

* p < .05

** p < .01

*** p < .001

Table 2

Demographics and Life Circumstances

Step 1	Predictor	Mean (SD)	<i>b</i>	
		%		
	Age	39.94 (11.55)	0.09 (0.05)	†
	Caucasian	83%	-1.17 (1.49)	
	Education (yrs)	14.59 (2.43)	0.33 (0.30)	
	Income	33.6 (33.48)	0.005 (0.02)	
	Married/cohabiting	34%	-1.25 (1.48)	
	Sex (female)	58%	-1.69 (1.13)	‡
	Unemployed	17%	2.59 (1.61)	‡
	Life events	7.01 (5.29)	-0.14 (0.13)	
	Intelligence	108.58 (10.57)	-0.13 (0.07)	*
	Age × Treatment		-0.14 (0.11)	‡
	Caucasian × Treatment		-4.82 (2.98)	‡
	Education (yrs) × Treatment		0.63 (0.60)	
	Income × Treatment		0.03 (0.04)	
	Marital Status × Treatment		-7.34 (2.95)	*
	Sex × Treatment		-0.13 (2.27)	
	Unemployment Status × Treatment		-8.92 (3.16)	**
	Life events × Treatment		-0.60 (0.26)	*
	Intelligence × Treatment		0.004 (0.13)	
Steps 2 & 3 (retain effects at $p < .20$, $p < .10$)				
	Age		0.12 (0.05)	*
	Marital Status		-1.38 (1.23)	
	Sex		-1.91 (1.05)	†
	Unemployment Status		2.60 (1.50)	†
	Life events		-0.09 (0.13)	
	Intelligence		-0.11 (0.05)	*
	Marital Status × Treatment		-7.75 (2.33)	**
	Unemployment Status × Treatment		-8.82 (2.98)	**
	Life events × Treatment		-0.52 (0.24)	*
Step 4 (retain effects at $p < .05$)				
<u>Non significant</u>				
	Marital Status		-1.41 (1.24)	
	Unemployment Status		2.82 (1.51)	†
	Life events		-0.10 (0.13)	
<u>Significant</u>				
	Age		0.13 (0.05)	*

Step 1	Predictor	Mean (SD)	<i>b</i>
		%	
	Intelligence		-0.10 (0.05) *
	Marital Status × Treatment		-7.35 (2.34) **
	Unemployment Status × Treatment		-9.11 (3.00) **
	Life events × Treatment		-0.53 (0.25) *

Note. yrs = years. Because the non-significant terms in Step 4 are constituents of the interaction terms, all items from Step 4 will be entered into the final prediction model.

[‡] p < .20

[†] p < .10

* p < .05

** p < .01

*** p < .001

Table 3

Final Model with all Significant Predictors from the Domain Models

Predictor	Mean (SD)	%	β^a	t (df)	p
Chronic MDD	51%		3.45 (1.02)	3.37 (158)	0.001
Age	39.94 (11.55)		0.10 (0.05)	2.00 (159)	0.047
Intelligence	108.58 (10.57)		-0.10 (0.05)	-2.07 (161)	0.04
Married/cohabiting	34%		-1.61 (1.20)	-1.34 (155)	0.18
Unemployed	17%		2.82 (1.46)	1.94 (161)	0.055
Life events	7.01 (5.29)		-0.07 (0.12)	-0.54 (157)	0.59
Marital Status \times Treatment			-8.18 (2.28)	-3.59 (157)	< 0.001
Unemployment Status \times Treatment			-9.61 (2.91)	-3.31 (162)	0.001
Life events \times Treatment			-0.49 (0.24)	-2.05 (158)	0.04

^a Note. Values represent unstandardized beta coefficients predicting HRSD scores at the end of acute treatment (Week 16 Scores). For lower-order terms, values represent effects averaged across the two treatment groups. Negative values represent more desirable outcomes (i.e., lower estimated week 16 scores). For interaction terms, values represent the difference between the two treatments in the magnitude of the effect. Parenthetical entries represent standard error estimates.

Table 4

Week 16 Score Estimates for Each of the Three Prognostic Variables

Prognostic Variable	Value	Week 16 HRSD
Age	-1 SD (28 yrs)	7.8 (1.01)
	Mean (40 yrs)	8.9 (0.76)
	+1 SD (51 yrs)	10.0 (0.88)
Intelligence	-1 SD (98)	9.9 (0.92)
	Mean (109)	8.9 (0.76)
	+1 SD (119)	7.8 (0.91)
Chronic MDD	Non-chronic	7.1 (0.94)
	Chronic	10.6 (0.9)

Note. Values represent estimated HRSD scores at the end of treatment, Week 16. Parenthetical entries represent standard error estimates.