



Published in final edited form as:

*Psychol Med.* 2015 November ; 45(15): 3205–3215. doi:10.1017/S0033291715001208.

## Predictors of outcome for telephone and face-to-face administered cognitive behavioral therapy for depression

C. Stiles-Shields, M. E. Corden, M. J. Kwasny, S. M. Schueller, and D. C. Mohr\*

Department of Preventive Medicine and Center for Behavioral Intervention Technologies, Northwestern University Feinberg School of Medicine, Chicago, IL, USA

### Abstract

**Background**—Cognitive behavioral therapy (CBT) can be delivered efficaciously through various modalities, including telephone (T-CBT) and face-to-face (FtF-CBT). The purpose of this study was to explore predictors of outcome in T-CBT and FtF-CBT for depression.

**Method**—A total of 325 depressed participants were randomized to receive eighteen 45-min sessions of T-CBT or FtF-CBT. Depression severity was measured using the Hamilton Depression Rating Scale (HAMD) and the Patient Health Questionnaire-9 (PHQ-9). Classification and regression tree (CART) analyses were conducted with baseline participant demographics and psychological characteristics predicting depression outcomes, HAMD and PHQ-9, at end of treatment (week 18).

**Results**—The demographic and psychological characteristics accurately identified 85.3% and 85.0% of treatment responders and 85.7% and 85.0% of treatment non-responders on the HAMD and PHQ-9, respectively. The Coping self-efficacy (CSE) scale predicted outcome on both the HAMD and PHQ-9; those with moderate to high CSE were likely to respond with no other variable influencing that prediction. Among those with low CSE, depression severity influenced response. Social support, physical functioning, and employment emerged as predictors only for the HAMD, and sex predicted response on the PHQ-9. Treatment delivery method (i.e. telephone or face-to-face) did not impact the prediction of outcome.

**Conclusions**—Findings suggest that the predictors of improved depression are similar across treatment modalities. Most importantly, a moderate to high level of CSE significantly increases the chance of responding in both T-CBT and FtF-CBT. Among patients with low CSE, those with lower depressive symptom severity are more likely to do well in treatment.

### Keywords

CART; CBT; depression; outcome; predictors; telephone therapy

---

\*Address for correspondence: D. C. Mohr, Ph.D., Department of Preventive Medicine and Center for Behavioral Intervention Technologies, Northwestern University Feinberg School of Medicine, Chicago, IL 60611, USA. (d-mohr@northwestern.edu).

### Declaration of Interest

None.

## Introduction

Depressive disorders are common mental health afflictions. In the United States alone, lifetime morbid risk of major depressive disorder is 29.9% and the 12-month prevalence is estimated to be between 6.6% and 10.3% (Kessler *et al.* 2003, 2005, 2012; Reeves *et al.* 2011). Psychological treatments for depression are effective (Cuijpers *et al.* 2008) and desirable to patients (Priest *et al.* 1996; Brody *et al.* 1997; Bedi *et al.* 2000; Churchill *et al.* 2000; Dwight-Johnson *et al.* 2000). Cognitive behavioral therapy (CBT) is the most studied psychotherapy for the treatment of depression and carries the strongest body of evidence for its effectiveness (Dobson, 1989; Butler *et al.* 2006; Cuijpers *et al.* 2013). Increasingly, the delivery of CBT through different media has been investigated, including via telephone (Mohr *et al.* 2005), internet websites (Andersson & Cuijpers, 2009), and bibliotherapy (Cuijpers, 1997). While CBT is effective when delivered across many different media, improvement is not inevitable in any of them.

A vast literature of baseline predictors of outcome for CBT for depression exists. Most predictors tend to be prognostic indicators, or factors that identify patients that may do better in treatment in general. Positive outcomes are often found among patients with: (1) lower symptom severity (Jarrett *et al.* 1991; Sotsky *et al.* 1991; Shapiro *et al.* 1994; Thase *et al.* 1994; Agosti & Ocepek-Welikson, 1997; Persons *et al.* 1999; Hamilton & Dobson, 2002; Coffman *et al.* 2007); (2) shorter current episode duration (Sotsky *et al.* 1991; Rush *et al.* 1977; Hamilton & Dobson, 2002); (3) absence of family history of depression (Sotsky *et al.* 1991); (4) older age of initial onset (Jarrett *et al.* 1991; Sotsky *et al.* 1991; Agosti & Ocepek-Welikson, 1997; Hamilton & Dobson, 2002); (5) lower number of previous episodes (Sotsky *et al.* 1991; Hamilton & Dobson, 2002; Bockting *et al.* 2006; Fournier *et al.* 2009); and (6) absence of comorbid conditions (Reich *et al.* 1995; Gelhart & King, 2001; Driessen & Hollon, 2010).

Identifying prognostic variables is difficult and applying them to clinical practice introduces additional challenges. Statistical techniques for identifying moderators are often underpowered to uncover these relationships in individual studies (Brown *et al.* 2013). Furthermore, most studies explore only a few variables that investigators selected *a priori*. These findings can then only support that the variable selected was meaningful, but not whether another variable would have been more useful. Practically, clinicians need to know which are the main factors to consider in determining questions of prognosis. However, attempting to predict outcomes from multiple predictors might lead to classification issues especially in cases of high multi-collinearity between predictors. Classification and Regression Tree (CART) analysis is a data-mining technique that uses recursive binary partitioning to select optimal splits of predictor variables to obtain increasingly homogenous groups with respect to the outcome (Breiman *et al.* 1984; King & Resick, 2014). The benefits of this method over more commonly used analyses for predicting outcomes include the interpretability (i.e. the trees demonstrate if-then conditions that are often considered to be easily interpreted), the capability of including many possible predictor variables, and that these models are robust to violations of normality and linearity to which typical regression models are bound. Thus, CART is an ideal method for exploratory analyses using a broad number of variables.

In exploring predictors of treatment response to CBT, it is important to understand them in the context of different forms of delivery beyond face-to-face (FtF-CBT). The most prominent alternative delivery medium for CBT is via the telephone (T-CBT). For the treatment of depression, T-CBT and FtF-CBT produce similar changes in depression severity at treatment completion (Mohr *et al.* 2012), and produce similar therapeutic alliance ratings from both the patient and therapist perspectives (Stiles-Shields *et al.* 2014b). As the use of the telephone to provide psychotherapy becomes more widespread (Novotney, 2011), it will be important to understand whether any subset of patients might be more likely to benefit from a particular approach. A number of studies have examined the usefulness of patient variables, such as marital status, employment status, severity of baseline depression, presence of personality disorders, stress, reactance, and internalizing *v.* externalizing coping to differentially predict responsiveness to various forms of psychotherapy and pharmacotherapy (Beutler *et al.* 1991; Dimidjian *et al.* 2006; Fournier *et al.* 2008). In addition, people with a diagnosis of co-morbid anxiety at baseline experience less benefit from T-CBT, relative to FtF-CBT (Stiles-Shields *et al.* 2014a). Thus, it is worth exploring whether predictors of outcome are similar across T-CBT and FtF-CBT.

The primary aim of this study was to use CART analyses to explore patient predictors of response to CBT for depression in T-CBT and FtF-CBT. An exploratory aim of the study was to explore whether patient predictors vary between treatment delivery method (*i.e.* telephone *v.* face-to-face).

## Method

This study is a secondary analysis of data from a randomized controlled trial comparing the efficacy and retention rates of T-CBT and FtF-CBT in a cohort of 325 depressed participants (Mohr *et al.* 2012).

## Participants

Recruitment of participants occurred from November 2007 to December 2010 from primary-care clinics located in an academic medical center in the Chicago area.

Participants were eligible for randomization if they met criteria for major depressive disorder, had a minimum score of 16 on the Hamilton Depression Rating Scale (HAMD), were at least 18 years of age, spoke English, and were able to participate in face-to-face or telephone therapy. Exclusion criteria included having visual or hearing impairments preventing participation, meeting criteria for depression of an organic etiology or a severe psychiatric disorder, reporting severe alcohol or substance abuse, meeting criteria for dementia; exhibiting severe suicidality (*i.e.* plan and intent), receiving or planning to receive individual psychotherapy, or initiation of antidepressant pharmacotherapy in the previous 10 days.

In compliance with the University's Institutional Review Board, participants were sent a consent form. Research staff reviewed the consent with them over the phone and participants were given an opportunity to ask questions. Consent forms were signed and returned prior to baseline interviews.

## Treatments

Participants were randomized to either T-CBT or FtF-CBT, stratified by antidepressant status and research study therapist ( $n = 9$ ) by a blinded statistician. The treatment delivery medium was the only experimental factor to vary between the two groups, with both treatments using the same CBT protocol (Beck, 1995) adapted and validated for use over the phone (Mohr *et al.* 2005). To eliminate therapist effects, PhD-level psychologists acted as therapists for both conditions. All therapists received training and supervision from the Beck Institute for Cognitive Behavioral Therapy. All therapy sessions were recorded and 8% were randomly selected and rated by the supervisor on the Cognitive Therapy Rating Scale (Vallis *et al.* 1986) for fidelity. Further detail of therapist training and fidelity are noted elsewhere (Mohr *et al.* 2012).

FtF-CBT participants were seen in the Preventive Medicine Clinic at Northwestern University, whereas T-CBT was conducted exclusively via the telephone. Participants in the T-CBT condition received instructions to conduct the telephone sessions in a private, safe, and distraction-free environment. All participants received eighteen 45-min sessions, with two sessions weekly for the first weeks, followed by 12 weekly sessions, and two final booster sessions over 4 weeks. Participants also received a client workbook that explained CBT concepts and provided worksheets for topics including behavioral activation, cognitive restructuring, and social support. Optional chapters addressed common co-morbidities, such as anxiety management, relaxation training, assertiveness training, anger management, and insomnia.

## Assessment

CART analyses included measures of depression severity, measured from baseline and end of treatment (week 18). Self-reported depression severity was measured using the Patient Health Questionnaire-9 (PHQ-9; Kroenke & Spitzer, 2002), which has high internal consistency (Cronbach's alphas were 0.75–0.91 for this trial) and face validity (Corson *et al.* 2004). Interviewer-based depression severity was evaluated using the 17-item HAMD (Hamilton, 1960). Bachelor-level research assistants, who were trained and supervised by a licensed clinical psychologist, administered the HAMD. To ensure inter-rater reliability, one audio-tape of the HAMD assessment was randomly selected every 1–2 weeks for calibration ratings with all evaluators. The mean interclass correlations were 0.96.

In addition to measures of depression severity, bachelor-level research assistants administered the Mini International Neuropsychiatric Interview (MINI; Sheehan *et al.* 1997) at baseline over the telephone. This semi-structured diagnostic interview diagnosed any DSM-IV-TR co-morbid conditions (APA, 2000). Baseline data included in the CART analysis included age, sex, ethnicity, race, marital status, education, employment status, household income, antidepressant medication status, trauma and abuse history (Wolfe & Kimberling, 1997), presence of a co-morbid anxiety disorder from the MINI (Sheehan *et al.* 1997; Stiles-Shields *et al.* 2014a), the HAMD total score, the PHQ-9 total score, the Insomnia Severity Index total score (ISI; Doghramji, 2006), the Life Experiences Survey total score (LES; Sarason *et al.* 1978), the Generalized Anxiety Disorder-7 total score (GAD-7; Spitzer *et al.* 2006), the Alcohol Use Disorders Identification Test total score

(AUDIT; Babor *et al.* 1992), the Positive and Negative Affect Scale total score (PANAS; Watson *et al.* 1988), Medical Outcomes Study 36 total and subscale (Vitality, Physical Functioning, Bodily Pain, General Health Perceptions, Physical Role Functioning, Emotional Role Functioning, Social Role Functioning, Mental Health) scores (SF-36; Brazier *et al.* 1992), the Brief Symptom Inventory total and subscale (Somatization, Obsessive-Compulsive, Interpersonal Sensitivity, Depression, Anxiety, Anger-Hostility, Phobic Anxiety, Paranoid Ideation, Psychoticism) scores (Lehman *et al.* 2012), the Coping Self-Efficacy Scale total score (CSE; Chesney *et al.* 2006), the Life Stressors and Social Resources total score (LISRES-A; Moos *et al.* 1988), the Nijmegen Motivation Questionnaire-2 total score (NML-2; Keijsers *et al.* 1999), Outcome Expectations Questionnaire – Patient Version, Perceived Barriers to Psychotherapy total score (PBP; Mohr *et al.* 2006), Social Provisions Scale total score (SPS; Russell & Cutrona, 1984), Scale for Interpersonal Behavior total score (SIBS; Arrindell & van der Ende, 1985), Perceived Stress Scale total score (PSS; Cohen *et al.* 1983), and Apathy Evaluation Scale total score (AES; Marin *et al.* 1991). Treatment assignment was also included to explore whether predictors would vary as a function of treatment delivery medium.

### Data analysis

CART analyses use recursive partitioning algorithms to find optimal ‘splits’ of any variable to separate responders from non-responders. For each systematic split or step, one or more rules are assessed to determine how or whether to proceed down the tree. This process results in a visualization that is reminiscent of an inverted tree, with a single root at the top that leads to branches, finalizing in non-branching leaves at the bottom. Each split or stop in a branch is referred to as a node. Each node denotes a predictor variable critical for that decision point and provides the amount of the sample that were impacted by that variable for the prediction (for a more detailed description of CART analyses in psychological treatment research, see King & Resick, 2014).

CART analyses were conducted with all baseline covariates. Treatment response was defined as end of treatment scores below 16 and 10 for the HAMD and PHQ-9, respectively. The cut-off of 16 on the HAMD was determined as it produces a comparative sample to that of a PHQ-9 cut-off of 10, which is consistent with the MacArthur recommendations for referrals to psychotherapy at the cut-off for mild depressive symptoms (The MacArthur Foundation’s Initiative on Depression and Primary Care, 2004). The cut-offs for both measures create a sample of roughly 2/3 response, which is consistent with response rates for the treatment of depression with CBT (Driessen & Hollon, 2010).

The trees were fit using the *rpart* package in R version 3.0.1 (R Core Team, 2013), and pruned with a complexity parameter set at the largest value that was within a standard error of the minimum cross-validated error.

For increased understanding and validation of the findings from the CART analyses, Random Forest analyses were also run (Breiman, 2001). Random Forests is a method of generating many trees and aggregating their results. This is done through creating a large number of trees, each of which is constructed using a bootstrap of the sample dataset. Each node in these trees is split using the best among a subset of predictors randomly chosen at

that node. An estimate of the error rate for Random Forests is obtained by predicting the data not in the bootstrap sample at each bootstrap iteration [referred to as ‘out of the bag’ (OOB) data]. The OOB data is beneficial in determining variable importance. The value of variable importance is measured through the program examining how much the prediction error (OOB) increases when data for that variable is permuted, with all other variables left unchanged (Breiman, 2001; Liaw & Wiener, 2002). Random Forests were run using the randomForest package in R version 3.0.1 (R Core Team, 2013) to obtain the OOB and variable importance for the HAMD and PHQ-9 prediction trees.

## Results

### Participants

Baseline participant demographic and clinical characteristics are displayed in Table 1. Among the 325 participants entered into the trial, there were no significant differences in demographics across treatment groups.

### Predictors of treatment response and non-response based on the HAMD

The treatment response rate based on HAMD < 16 was 66.2% (49.5% for T-CBT and 50.5% for FtF-CBT). The CART model, pruned using the prune command, generated a tree based on the outcome of dichotomized HAMD scores at end of treatment. Fig. 1 displays the pruned tree. The model predicted that 231 participants would be treatment responders; 197 (85.3%) of these were accurately predicted. Of the 56 participants who were predicted to be treatment non-responders 48 (85.7%) were accurately predicted.

Variables (and scores) that predicted response included CSE ( $\geq 104$ ), baseline depression severity (HAMD < 23, PHQ-9 < 17), physical functioning (PF  $\geq 27$ ), social support (SPS < 65), education level [some high school, general education diploma (GED), or bachelor’s/ master’s degree], and employment status (employed or unemployed). Variables (and scores) that predicted non-response included CSE ( $< 104$ ), baseline depression severity (HAMD  $\geq 23$ , PHQ-9  $\geq 17$ ), PF ( $< 27$ ), social support (SPS  $\geq 65$ ), education level (some college or professional degree), and employment status (disability or retired). No other variables, including treatment assignment to FtF-CBT or T-CBT were related to response.

### Predictors of treatment response and non-response based on the PHQ-9

The treatment response rate based on PHQ-9 < 10 was 73.8% (50.2% for T-CBT and 49.8% for FtF-CBT). The CART model, pruned using the prune command, generated a tree based on the outcome of dichotomized PHQ-9 scores at end of treatment. Fig. 2 displays the pruned tree. The model predicted 240 participants would be treatment responders; 204 (85.0%) of these were accurately predicted. Of the 46 participants predicted to be treatment non-responders, 39 (85.0%) were accurately predicted. Variables (and scores) that predicted response included CSE ( $\geq 104$ ), baseline depression severity (PHQ-9 < 17, HAMD < 26), education level (some college, bachelor’s degree, or professional degree), and being male. Variables (and scores) that predicted non-response included CSE ( $< 104$ ), baseline depression severity (PHQ-9  $\geq 17$ , HAMD  $\geq 26$ ), education level (some high school, GED or

master's degree/PhD), and being female. No other variables, including treatment assignment to FtF-CBT or T-CBT were related to response.

### Variables of importance

The Random Forest analyses yielded an OOB error rate of 30.8% for both the HAMD and PHQ-9 models. The top values of variable importance for both the HAMD and PHQ-9 are presented in Table 2. The top variables of importance for the HAMD Random Forests were education, CSE, interviewer-based assessment of depression through the HAMD, age, and self-reported anxiety through the GAD-7. The top variables of importance for the PHQ-9 Random Forests were CSE, self-report depression through the PHQ-9, interviewer-based assessment of depression through the HAMD, education, and self-reported anxiety through the GAD-7.

### Discussion

Demographic and psychological characteristics at baseline of participants receiving outpatient T-CBT and FtF-CBT for depression accurately identified 85.3% and 85.0% of treatment responders and 85.7% and 85.0% of treatment non-responders with response defined by the HAMD and PHQ-9, respectively. CSE, baseline depression severity, and education consistently predicted outcome using both the HAMD and PHQ-9. Social support, PF, and employment emerged as predictors only for the HAMD, and sex predicted response on the PHQ-9. Treatment delivery method (i.e. telephone or face-to-face) and presence of co-morbid anxiety did not impact the prediction of outcome. These findings were supported through the Random Forests top variables of importance, including CSE, baseline depression, and education.

Participant baseline scores on CSE, which measures a person's confidence in his or her ability to cope effectively with situations that are appraised to be stressful (Chesney *et al.* 2006), was the primary predictive branch of the CART trees for response to treatment, regardless of how depression was measured. This measure alone predicted positive response outcomes for 41.3% and 40.2% of the total sample for the HAMD and PHQ-9, respectively. The value identified, surprisingly equivalent for both the HAMD and PHQ-9 (CSE = 104), is indicative of a moderate level of CSE according to previously established norms (Chesney *et al.* 2006), meaning that those with moderate levels of CSE or higher are more likely to respond relative to those with lower levels. To our knowledge, CSE has not been directly investigated or reported as a predictor of outcome in CBT for depression. However, studies investigating related concepts have shown findings consistent with these. For example, resourcefulness (Simons *et al.* 1985) as well as increased stressful life events (Fournier *et al.* 2009) have both been found to be prescriptive predictors of which patients benefit more from CBT compared to anti-depressant medication.

The CSE taps a person's confidence in his or her ability to cope effectively with situations that are appraised to be stressful (Chesney *et al.* 2006). CBT often requires patients to confront difficult or stressful thoughts or situations through cognitive restructuring, exposure, behavioral experiments, and other strategies (Beck, 1995; Simos & Hofmann, 2013). Patients who come into therapy with high levels of CSE may be more willing and

capable of tolerating the distress created by these strategies and more likely in general to persevere and succeed. This suggests that patients with low CSE may achieve better results if greater focus on enhancing CSE is provided earlier in treatment, which may provide patients more confidence in discussing and facing difficult life situations. Further research is needed to cross-validate these findings and explore their treatment implications.

It was interesting that the CART analyses found for both the PHQ-9 and HAMD outcomes, moderate to high CSE was the optimal split predicting response, with no additional covariates necessary. For those with low CSE, two constructs provided additional predictive value consistently across both models: depression and education. The finding that education consistently predicts response is not consistent with the available evidence indicating that education is not predictive of outcome in CBT (Jarrett *et al.* 1991; Hamilton & Dobson, 2002). Indeed, no gradient was identified in the current analyses; having attained a master's degree was grouped with some high school, GED, and college; while some college education and professional degrees are grouped in another. Furthermore, some of the educational categories (e.g. some college, professional degree) predict response for one depression outcome, and non-response in the other.

By contrast, the finding that baseline depressive symptom severity contributes to the prediction of treatment response is consistent with and extends previous findings (Jarrett *et al.* 1991; Sotsky *et al.* 1991; Shapiro *et al.* 1994; Thase *et al.* 1994; Agosti & Ocepek-Welikson, 1997; Persons *et al.* 1999; Hamilton & Dobson, 2002; Coffman *et al.* 2007). However, severity of depressive symptoms was only valuable as a predictor of treatment response among those with CSE, suggesting that greater severity of symptoms only exerts its negative influence among those who do not have the confidence in the coping skills to manage stress and distress. This finding points to the utility of CART analyses in uncovering potentially complex relationships among predictor variables.

Other variables provided predictive value, however not consistently across the models. In the HAMD model, among those with lower CSE, more depressed individuals and those with higher perceived social support were at most risk for non-response. This finding is inconsistent with the literature on the influence of social support on outcomes for treatment of depression (George *et al.* 1989); however, given that social support was not a top variable of importance in the Random Forests, we are hesitant to over-interpret these findings.

Among individuals with low CSE and depressive severity, poorer PF predicted non-response in the HAMD model. Among those with both low CSE and higher depressive severity, individuals with good social support and employment further differentiated responders from non-responders. However, the grouping of employment (disabled and retired *v.* employed and unemployed) suggested that the variable might be acting as a surrogate for PF. Sex entered only the PHQ-9 model, with women with low CSE and higher depressive severity at greater risk of non-response than their male counterparts. However, the inconsistency of these variables across the models suggests that the findings may not be reliable.

Randomization to T-CBT or FtF-CBT did not appear in the CART analyses as a predictor variable. This is consistent with the findings of the parent trial, finding no difference in



depression severity at post-treatment across the two treatment arms (Mohr *et al.* 2012). Thus, this investigator found no support for any prescriptive variables, it appears as though characteristics that predict response held regardless of the modality used. This might be due to the fact that the treatments were quite similar and the only difference in modality was whether the patient was in the room with the therapist or over the phone. Prescriptive variables might be more important for treatments that differ more markedly in mechanism of action such as psychotherapy *v.* medication. However, for a given type of treatment (e.g. CBT), prescriptive variables might be more valuable for media that differ more markedly, such as face-to-face *v.* internet CBT.

A number of other predictor variables included in the CART analyses that the literature has identified as predictors of CBT outcome, such as stress, motivation for treatment, and comorbid anxiety, were included in the analyses, yet they were not significant predictors for this sample. It may be that CSE and depressive symptom severity account for most of the variance, leaving little room for other variables, which might, on their own, have some predictive power. Specifically, CSE may be associated with traits such as anxiety, and studies examining anxiety may not have accounted for coping, as it has not previously been identified as a potential confounder. This demonstrates the potential strength of CART analyses to identify new relationships one might not have predicted that could be more important than previously established relationships. It also demonstrates the clinical utility of CART analyses, as the interaction options provided by trees maps more closely to intuitive clinical decision making and is a further step towards personalized medicine.

There are several limitations and caveats that should be considered in interpreting these results. First, as a secondary, exploratory analysis, these findings should be viewed with caution until they are replicated or refuted. Second, this trial examined CBT for depression; it is unclear how these results generalize to other forms of psychotherapy, other mental health conditions, and other treatment delivery media. Third, while the sample was ethnically diverse, participants were fairly well educated. The small number of participants with lower levels of education may be partly responsible for the inconsistent findings with respect to education and certainly limit generalizability to less-educated populations. Fourth, we used response criteria based on level of symptoms used for referral or initiation of treatment. For the HAMD, this was identical to entry criteria. While patients overall showed a strong response to treatment, it is possible that a few patients may only have moved 1 or 2 points on the HAMD to reach response criteria. Furthermore, these decision trees may not generalize to other criteria, such as full remission of symptoms. Finally, although CART analysis is an effective data-mining tool, there are some disadvantages to these models. Trees minimize total variability in the data, both from the population as well as sample variability. Without adequate pruning, trees are often overfit, in that they minimize sample variability. Additionally, misclassification errors may build. Since each branch depends on a previous one, misclassification errors in early branches will continue down the tree. CART is a form of data mining, thus, these findings should be validated in another dataset before generalizing these findings more broadly. Nevertheless, a strength of CART analyses is that they can identify new relationships one might not have predicted that could be more important than previously established relationships. Additionally, a sensitivity analysis using

Random Forests supported the findings of the CART analyses. Thus, although CART is an exploratory method, it has the potential to identify relationships that can be investigated in subsequent studies.

To the best of our knowledge, this is the first study to use exploratory CART analyses to evaluate patient characteristics that predict depression outcomes among patients receiving T-CBT or FtF-CBT for the treatment of major depressive disorder. The findings of the present study indicate that depressed patients with moderate to high CSE are likely to do well with CBT, regardless of other baseline characteristics. Among those with poor CSE, lower levels of depressive symptom severity also indicate a likely positive response. Those with low CSE and high depressive symptom severity were consistently found to be the most at risk of non-response. There was also a suggestion that in this group of low CSE, depressed individuals with low social support may be at risk for non-response. While these findings should be confirmed in future research, they point to the possibility of improving outcomes by enhancing CSE early in treatment as a strategy of mitigating potential treatment non-response to improve the overall impact of CBT.

## Acknowledgments

This research was supported by grants from the National Institute of Mental Health R01 MH059708 and R01 MH095753 (PI: Mohr); K08 MH102336 (PI: Schueller); and F31 MH106321 (PI: Stiles-Shields).

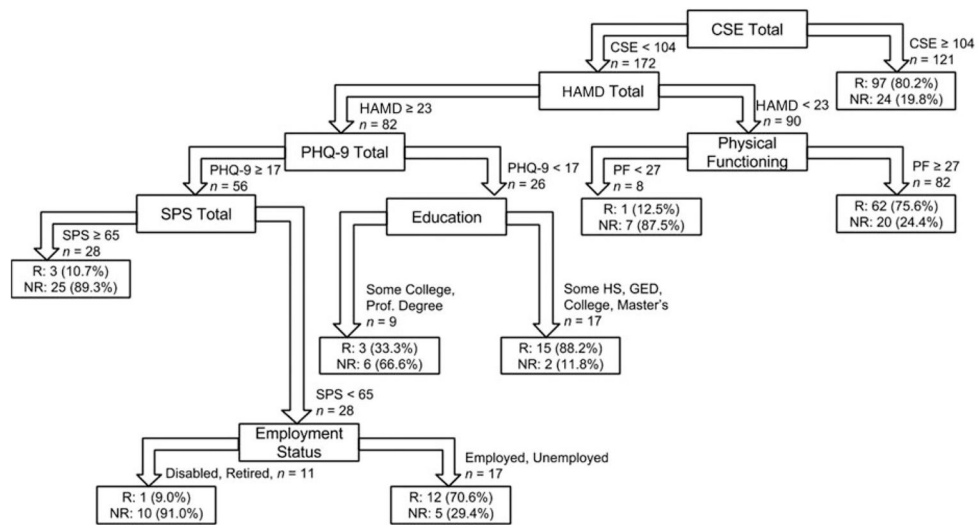
## References

- 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]
- American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 4. American Psychiatric Association; Washington, DC: 2000. Text Revision
- Andersson G, Cuijpers P. Internet-based and other computerized psychological treatments for adult depression: a meta-analysis. *Cognitive Behavior Therapy*. 2009; 38:196–205.
- Arrindell W, Van Der Ende J. Cross-sample invariance of the structure of self-reported distress and difficulty in assertiveness. *Advances in Behaviour Research and Therapy*. 1985; 7:205–243.
- Babor, TF.; De La Fuente, JR.; Saunders, J.; Grant, M. *The Alcohol Use Disorders Identification Test: Guidelines for Use in Primary Health Care*. World Health Organization; Geneva: 1992. (WHO Publication No. 92.4)
- Beck, JS. *Cognitive Therapy*. Guilford Press; New York: 1995.
- Bedi N, Chilvers C, Churchill R, Dewey M, Duggan C, Fielding K, Gretton V, Miller P, Harrison G, Lee A, Williams I. Assessing effectiveness of treatment of depression in primary care. Partially randomised preference trial. *British Journal of Psychiatry*. 2000; 177:312–318. [PubMed: 11116771]
- Beutler LE, Engle D, Mohr D, Daldrup RJ, Bergan J, Meredith K, Merry W. Predictors of differential response to cognitive, experiential, and self-directed psychotherapeutic procedures. *Journal of Consulting and Clinical Psychology*. 1991; 59:333–340. [PubMed: 2030196]
- Bockting CL, Spinhoven P, Koeter MW, Wouters LF, Visser I, Schene AH, Group DS. Differential predictors of response to preventive cognitive therapy in recurrent depression: a 2-year prospective study. *Psychotherapy and Psychosomatics*. 2006; 75:229–236. [PubMed: 16785772]
- Brazier JE, Harper R, Jones NM, O'cathain A, Thomas KJ, Usherwood T, Westlake L. Validating the SF-36 health survey questionnaire: new outcome measure for primary care. *British Medical Journal*. 1992; 305:160–164. [PubMed: 1285753]

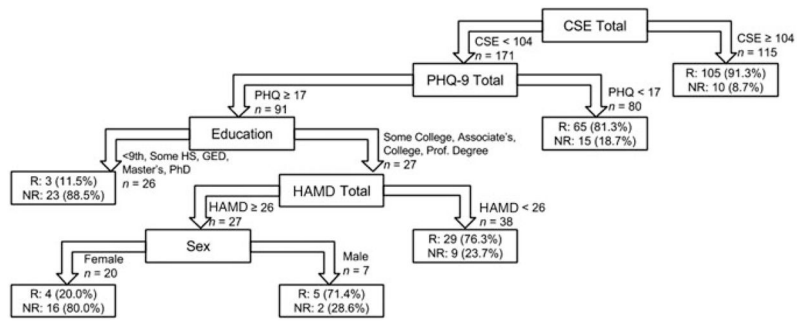
- Breiman L. Random forests. *Machine Learning*. 2001; 45:5–32.
- Breiman, L.; Friedman, J.; Olshen, RA.; Stone, CJ. *Classification and Regression Trees*. CRC Press; New York: 1984.
- Brody DS, Khaliq AA, Thompson TL II. Patients' perspectives on the management of emotional distress in primary care settings. *Journal of General Internal Medicine*. 1997; 12:403–406. [PubMed: 9229277]
- Brown CH, Sloboda Z, Faggiano F, Teasdale B, Keller F, Burkhart G, Vigna-Taglianti F, Howe G, Masyn K, Wang W. Methods for synthesizing findings on moderation effects across multiple randomized trials. *Prevention Science*. 2013; 14:144–156. [PubMed: 21360061]
- Butler AC, Chapman JE, Forman EM, Beck AT. The empirical status of cognitive-behavioral therapy: a review of meta-analyses. *Clinical Psychology Review*. 2006; 26:17–31. [PubMed: 16199119]
- Chesney MA, Neilands TB, Chambers DB, Taylor JM, Folkman S. A validity and reliability study of the coping self-efficacy scale. *British Journal of Health Psychology*. 2006; 11:421–437. [PubMed: 16870053]
- Churchill R, Khaira M, Gretton V, Chilvers C, Dewey M, Duggan C, Lee A. Treating depression in general practice: factors affecting patients' treatment preferences. *British Journal of General Practice*. 2000; 50:905–906. [PubMed: 11141877]
- Coffman SJ, Martell CR, Dimidjian S, Gallop R, Hollon SD. Extreme nonresponse in cognitive therapy: can behavioral activation succeed where cognitive therapy fails? *Journal of Consulting and Clinical Psychology*. 2007; 75:531–541. [PubMed: 17663608]
- Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. *Journal of Health and Social Behavior*. 1983; 24:385–396. [PubMed: 6668417]
- Corson K, Gerrity MS, Dobscha SK. Screening for depression and suicidality in a VA primary care setting: 2 items are better than 1 item. *American Journal of Managed Care*. 2004; 10:839–845. [PubMed: 15609737]
- Cuijpers P. Bibliotherapy in unipolar depression: a meta-analysis. *Journal of Behavior Therapy and Experimental Psychiatry*. 1997; 28:139–147. [PubMed: 9194011]
- Cuijpers P, Berking M, Andersson G, Quigley L, Kleiboer A, Dobson KS. A meta-analysis of cognitive-behavioural therapy for adult depression, alone and in comparison with other treatments. *Canadian Journal of Psychiatry*. 2013; 58:376–385.
- Cuijpers P, Van Straten A, Andersson G, Van Oppen P. Psychotherapy for depression in adults: a meta-analysis of comparative outcome studies. *Journal of Consulting and Clinical Psychology*. 2008; 76:909–922. [PubMed: 19045960]
- Dimidjian S, Hollon SD, Dobson KS, Schmalong KB, Kohlenberg RJ, Addis ME, Gallop R, Mcglinchey JB, Markley DK, Gollan JK, Atkins DC, Dunner DL, Jacobson NS. Randomized trial of behavioral activation, cognitive therapy, and antidepressant medication in the acute treatment of adults with major depression. *Journal of Consulting and Clinical Psychology*. 2006; 74:658–670. [PubMed: 16881773]
- Dobson KS. A meta-analysis of the efficacy of cognitive therapy for depression. *Journal of Consulting and Clinical Psychology*. 1989; 57:414–419. [PubMed: 2738214]
- Doghramji K. The epidemiology and diagnosis of insomnia. *American Journal of Managed Care*. 2006; 12:S214–S220. [PubMed: 16686591]
- Diessen E, Hollon SD. Cognitive behavioral therapy for mood disorders: efficacy, moderators and mediators. *Psychiatric Clinics of North America*. 2010; 33:537–555. [PubMed: 20599132]
- Dwight-Johnson M, Sherbourne CD, Liao D, Wells KB. Treatment preferences among depressed primary care patients. *Journal of General Internal Medicine*. 2000; 15:527–534. [PubMed: 10940143]
- 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]
- Fournier JC, Derubeis RJ, Shelton RC, Hollon SD, Amsterdam JD, Gallop R. Prediction of response to medication and cognitive therapy in the treatment of moderate to severe depression. *Journal of Consulting and Clinical Psychology*. 2009; 77:775–787. [PubMed: 19634969]

- Gelhart RP, King HL. The influence of comorbid risk factors on the effectiveness of cognitive-behavioral treatment of depression. *Cognitive and Behavioral Practice*. 2001; 8:18–28.
- George LK, Blazer DG, Hughes DC, Fowler N. Social support and the outcome of major depression. *British Journal of Psychiatry*. 1989; 154:478–485. [PubMed: 2590779]
- Hamilton KE, Dobson KS. Cognitive therapy of depression: pretreatment patient predictors of outcome. *Clinical Psychology Review*. 2002; 22:875–893. [PubMed: 12214329]
- Hamilton M. A rating scale for depression. *Journal of Neurology, Neurosurgery, and Psychiatry*. 1960; 23:56–62.
- Jarrett RB, Eaves GG, Grannemann BD, Rush AJ. Clinical, cognitive, and demographic predictors of response to cognitive therapy for depression: a preliminary report. *Psychiatry Research*. 1991; 37:245–260. [PubMed: 1891508]
- Keijsers PJ, Schaap PDR, Hoogduin C, Hoogsteyns B, De Kemp ECM. Preliminary results of a new instrument to assess patient motivation for the treatment of cognitive-behavioural therapy. *Behavioural and Cognitive Psychotherapy*. 1999; 27:165–179.
- Kessler RC, Berglund P, Demler O, Jin R, Koretz D, Merikangas KR, Rush AJ, Walters EE, Wang PS. The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). *Journal of the American Medical Association*. 2003; 289:3095–3105. [PubMed: 12813115]
- Kessler RC, Chiu WT, Demler O, Walters EE. Prevalence, severity, and comorbidity of twelve-month DSM-IV disorders in the National Comorbidity Survey Replication (NCS-R). *Archives of General Psychiatry*. 2005; 62:617–627. [PubMed: 15939839]
- Kessler RC, Petukhova M, Sampson NA, Zaslavsky AM, Wittchen HU. Twelve-month and lifetime prevalence and lifetime morbid risk of anxiety and mood disorders in the United States. *International Journal of Methods in Psychiatric Research*. 2012; 21:169–184. [PubMed: 22865617]
- King MW, Resick PA. Data mining in psychological treatment research: a primer on classification and regression trees. *Journal of Consulting and Clinical Psychology*. 2014; 82:895–905. [PubMed: 24588404]
- Kroenke K, Spitzer RL. The PHQ-9: a new depression diagnostic and severity measure. *Psychiatric Annals*. 2002; 32:1–7.
- Lehman KA, Burns MN, Gagen EC, Mohr DC. Development of the brief inventory of perceived stress. *Journal of Clinical Psychology*. 2012; 68:631–644. [PubMed: 22467381]
- Liaw A, Wiener M. Classification and regression by randomForest. *R News*. 2002; 2:18–22.
- Marin RS, Biedrzycki RC, Firinciogullari S. Reliability and validity of the Apathy Evaluation Scale. *Psychiatry Research*. 1991; 38:143–162. [PubMed: 1754629]
- Mohr DC, Hart SL, Howard I, Julian L, Vella L, Catledge C, Feldman MD. Barriers to psychotherapy among depressed and nondepressed primary care patients. *Annals of Behavioral Medicine*. 2006; 32:254–258. [PubMed: 17107299]
- Mohr DC, Hart SL, Julian L, Catledge C, Honos-Webb L, Vella L, Tasch ET. Telephone-administered psychotherapy for depression. *Archives of General Psychiatry*. 2005; 62:1007–1014. [PubMed: 16143732]
- Mohr DC, Ho J, Duffecy J, Reifler D, Sokol L, Burns MN, Jin L, Siddique J. Effect of telephone-administered vs. face-to-face Cognitive Behavioral Therapy on Adherence to therapy and depression outcomes among primary care patients: a randomized trial. *Journal of the American Medical Association*. 2012; 307:2278–2285. [PubMed: 22706833]
- Moos RH, Fenn CB, Billings AG. Life stressors and social resources: an integrated assessment approach. *Social Science and Medicine*. 1988; 27:999–1002. [PubMed: 3227394]
- Novotney A. A new emphasis on telehealth: how can psychologists stay ahead of the curve – and keep patients safe? *Monitor on Psychology*. 2011; 42:40.
- Persons JB, Bostrom A, Bertagnolli A. Results of randomized controlled trials of cognitive therapy for depression generalize to private practice. *Cognitive Therapy and Research*. 1999; 23:535–548.
- Priest RG, Vize C, Roberts A, Roberts M, Tylee A. Lay people's attitudes to treatment of depression: results of opinion poll for Defeat Depression Campaign just before its launch. *British Medical Journal*. 1996; 313:858–859. [PubMed: 8870574]

- R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2013.
- Reeves WC, Strine TW, Pratt LA, Thompson W, Ahluwalia I, Dhingra SS, McKnight-Eily LR, Harrison L, D'angelo DV, Williams L, Morrow B, Gould D, Safran MA. Centers for Disease Control and Prevention (CDC). Mental illness surveillance among adults in the United States. *Morbidity and Mortality Weekly Report. Surveillance Summaries*. 2011; 60 (Suppl 3):1–29. [PubMed: 21881550]
- Reich J, Warshaw M, Peterson LG, White K. Comorbidity of panic and major depressive disorder. *Journal of Psychiatric Research*. 1995; 20:23–33.
- Rush AJ, Beck AT, Kovacs M, Hollon SD. Comparative efficacy of cognitive therapy and pharmacotherapy in the treatment of depressed outpatients. *Cognitive Therapy and Research*. 1977; 1:17–37.
- Russell, DW.; Cutrona, CE. Social Provisions Scale. Paper presented at the American Psychological Association Meeting; August, 1984; Toronto, ON, Canada. Iowa State University; 1984.
- Sarason IG, Johnson JH, Siegel JM. Assessing the impact of life changes: development of the Life Experiences Survey. *Journal of Consulting and Clinical Psychology*. 1978; 46:932–946. [PubMed: 701572]
- Shapiro DA, Barkham M, Rees A, Hardy GE, Reynolds S, Startup M. Effects of treatment duration and severity of depression on the effectiveness of cognitive-behavioral and psychodynamic-interpersonal psychotherapy. *Journal of Consulting and Clinical Psychology*. 1994; 62:522–534. [PubMed: 8063978]
- Sheehan DV, Lecrubier Y, Sheehan KH, Janavs J, Weiller E, Keskiner A, Schinka E, Knapp E, Sheehan MF, Dunbar GC. The validity of the Mini International Neuropsychiatric Interview (MINI) according to the SCID-P and its reliability. *European Psychiatry*. 1997; 12:232–241.
- Simons AD, Lustman PJ, Wetzel RD, Murphy GE. Predicting response to cognitive therapy of depression: the role of learned resourcefulness. *Cognitive Therapy and Research*. 1985; 9:79–89.
- Simos, G.; Hofmann, SG. *CBT for Anxiety Disorders: A Practitioner Book*. John Wiley & Sons; Oxford: 2013.
- Sotsky SM, Glass DR, Shea MT, Pilkonis PA, Collins JF, Elkin I, Watkins JT, Imber SD, Leber WR, Moyer J, Oliveri ME. 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 RL, Kroenke K, Williams JB, Lowe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Archives of Internal Medicine*. 2006; 166:1092–1097. [PubMed: 16717171]
- Stiles-Shields C, Kwasny M, Cai X, Mohr DC. Comorbid anxiety as a differential treatment predictor for telephone vs. face-to-face administered cognitive behavioral therapy for depression. *Depression and Anxiety*. 2014a; 31:934–940. [PubMed: 25407583]
- Stiles-Shields C, Kwasny MJ, Cai X, Mohr DC. Therapeutic alliance in face-to-face and telephone-administered cognitive behavioral therapy. *Journal of Consulting and Clinical Psychology*. 2014b; 82:349–354. [PubMed: 24447003]
- Thase ME, Reynolds CF, Frank E, Simons AD, Garamoni GD, McGeary J, Harden T, Fasiczka AL, Cahalane JF. Response to cognitive-behavioral therapy in chronic depression. *Journal of Psychotherapy Practice and Research*. 1994; 3:204–214.
- The Macarthur Foundation's Initiative on Depression and Primary Care. *The Macarthur Initiative on Depression and Primary Care at Dartmouth and Duke: Depression Management Toolkit*. Dartmouth; Hanover, NH: 2004.
- Vallis TM, Shaw BF, Dobson KS. The cognitive therapy scale: psychometric properties. *Journal of Consulting and Clinical Psychology*. 1986; 54:381–385. [PubMed: 3722567]
- 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]
- Wolfe, J.; Kimberling, R. Gender issues in the assessment of posttraumatic stress disorder. In: Wilson, J.; Keane, TM., editors. *Assessing Psychological Trauma and PTSD*. Guilford; New York: 1997. p. 192-238.



**Fig. 1.** Classification and regression tree model developed for predicting outcomes on the Hamilton Depression Rating Scale (HAMD). CSE, Coping Self-Efficacy Scale; PHQ-9, Patient Health Questionnaire-9; PF, physical functioning (as measured by SF-36); SPS, Social Provisions Scale.



**Fig. 2.** Classification and regression tree model developed for predicting outcomes on the Patient Health Questionnaire-9 (PHQ-9). CSE, Coping Self-Efficacy Scale; HAMD, Hamilton Depression Rating Scale.

**Table 1**

Participants' baseline demographics and psychiatric characteristics

Variable	FtF-CBT ( <i>n</i> = 162)	T-CBT ( <i>n</i> = 163)	<i>p</i> value
Age, mean (S.D.)	47.5 (13.5)	47.8 (12.6)	0.87
Female, <i>n</i> (%)	127 (78.4)	125 (76.7)	0.71
Ethnicity, <i>n</i> (%)			
Hispanic or Latino	21 (13.0)	23 (14.2)	0.76
Race, <i>n</i> (%)			
African American	36 (24.0)	36 (24.3)	
Caucasian	98 (65.3)	89 (60.1)	0.63
More than one race	12 (8.0)	18 (12.2)	
Other	4 (2.7)	5 (3.4)	
Married/cohabitating, <i>n</i> (%)	51 (31.7)	56 (34.4)	0.61
Education, <i>n</i> (%)			
<9th grade	0 (0.0)	2 (1.2)	
Some high school	2 (1.2)	7 (4.3)	
General Education Diploma	12 (7.4)	13 (8.0)	
Some college	31 (19.1)	26 (16.0)	
Associate's degree	10 (6.2)	14 (8.6)	0.77
Bachelor's degree	64 (39.5)	55 (33.7)	
Master's degree	30 (18.5)	33 (20.2)	
Doctoral degree	4 (2.5)	5 (3.1)	
Professional degree	9 (5.6)	10 (6.1)	
Employment, <i>n</i> (%)			
Employed	90 (55.6)	95 (58.3)	
Unemployed	33 (20.4)	32 (19.6)	
Disability	10 (6.2)	16 (9.8)	0.30
Retired	18 (11.1)	16 (9.8)	
Other	11 (6.8)	4 (2.5)	
Household income, US\$, mean (S.D.)	73 480 (64 057)	75 637 (75 989)	0.79
On active dose of antidepressant medication, <i>n</i> (%)	56 (34.6)	54 (33.1)	0.78
Total number of co-morbid diagnoses, mean (S.D.)	1.34 (1.06)	1.3 (1.11)	0.67
Co-morbid anxiety, <i>n</i> (%)			
Agoraphobia, panic	11 (8.7)	12 (10.2)	0.70
GAD	114 (90.5)	106 (89.8)	0.87
OCD	13 (10.3)	9 (7.6)	0.46
PTSD	15 (11.9)	8 (6.8)	0.17
Social phobia	25 (19.8)	33 (28.0)	0.14
HAMD, mean (S.D.)	22.8 (4.6)	22.9 (4.6)	0.77
PHQ-9, mean (S.D.)	16.4 (4.8)	17.2 (4.7)	0.12
GAD-7, mean (S.D.)	12.3 (5.0)	11.8 (4.7)	0.33
PANAS, mean (S.D.)	19.8 (5.9)	19.9 (6.0)	0.89



Variable	FtF-CBT ( <i>n</i> = 162)	T-CBT ( <i>n</i> = 163)	<i>p</i> value
LES, mean (S.D.)	-12.0 (10.6)	-11.4 (11.5)	0.65
NML, mean (S.D.)	148.2 (13.1)	148.2 (13.3)	0.98
SPS, mean (S.D.)	70.2 (11.9)	70.6 (10.6)	0.79
CSE, mean (S.D.)	91.0 (42.9)	98.2 (43.1)	0.14
AUDIT, mean (S.D.)	3.7 (5.0)	3.6 (4.5)	0.94
PBP, mean (S.D.)	40.0 (11.3)	39.5 (10.7)	0.85
SIBS, mean (S.D.)	11.0 (5.8)	10.4 (5.5)	0.34
PSS, mean (S.D.)	27.1 (5.2)	26.8 (5.4)	0.50
AES, mean (S.D.)	48.4 (9.7)	49.8 (8.6)	0.14
SF-36, mean (S.D.)			
Bodily pain	57.9 (26.6)	51.2 (26.1)	0.03
General health perception	50.1 (22.2)	47.2 (24.4)	0.24
Mental health	33.6 (13.6)	34.0 (12.4)	0.45
Physical functioning	74.1 (27.4)	66.3 (30.7)	0.02
Emotional role functioning	37.1 (22.8)	38.7 (24.3)	0.47
Physical role functioning	62.1 (31.6)	57.7 (33.8)	0.25
Social role functioning	37.7 (23.3)	34.2 (20.4)	0.26
Vitality	22.4 (14.5)	21.0 (14.4)	0.34
Physical components summary	51.7 (11.8)	48.3 (12.6)	0.02
Mental components summary	24.5 (9.4)	25.4 (8.2)	0.17
Trauma/abuse history, <i>n</i> (%)	92 (57.5)	99 (59.9)	0.67

FtF-CBT, Face-to-Face cognitive behavioral therapy; T-CBT, telephone cognitive behavioral therapy; HAMD, Hamilton Depression Rating Scale; PHQ-9, Patient Health Questionnaire-9; GAD-7, Generalized Anxiety Disorder Questionnaire-7; PANAS, Positive and Negative Affect Schedule; LES, Life Experiences Survey; NML, Nijmegen Motivation; SPS, Social Provisions Scale; CSE, Coping Self-Efficacy Scale; AUDIT, Alcohol Use Disorders Identification Test; PBP, Perceived Barriers to Psychotherapy; SIBS, Scale for Interpersonal Behavior; PSS, Perceived Stress Scale; AES, Apathy Evaluation Scale; SF-36, Medical Outcomes Study 36.

**Table 2**

Top variables of importance as determined through Random Forest analyses

<b>HAMD</b>		<b>PHQ-9</b>	
<b>Variable</b>	<b>Importance value</b>	<b>Variable</b>	<b>Importance value</b>
Education	10	CSE	14
CSE	9	PHQ-9	11
HAMD	7	HAMD	10
Age	7	Education	8
GAD-7	7	GAD-7	6

HAMD, Hamilton Depression Rating Scale; PHQ-9, Patient Health Questionnaire-9; CSE, Coping Self-Efficacy Scale; GAD-7, Generalized Anxiety Disorder Questionnaire-7.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript