



Published in final edited form as:

J Consult Clin Psychol. 2013 August ; 81(4): . doi:10.1037/a0032879.

Nomothetic and Idiographic Symptom Change Trajectories in Acute-Phase Cognitive Therapy for Recurrent Depression

Jeffrey R. Vittengl,

Department of Psychology, Truman State University

Lee Anna Clark,

Department of Psychology, University of Notre Dame

Michael E. Thase, and

Department of Psychiatry, Perelman School of Medicine, University of Pennsylvania

Robin B. Jarrett

Department of Psychiatry, The University of Texas Southwestern Medical Center.

Abstract

Objective—We tested nomothetic and idiographic convergence and change in three symptom measures during acute-phase cognitive therapy (CT) for depression and compared outcomes among patients showing different change patterns.

Method—Outpatients ($N = 362$; 69% women; 85% white; age mean = 43 years) with *DSM-IV* recurrent major depressive disorder completed the Hamilton Rating Scale for Depression (Hamilton, 1960), Beck Depression Inventory (Beck, Ward, Mendelson, Mock, & Erbaugh 1961), and Inventory for Depressive Symptomatology—Self-Report (Rush, Gullion, Basco, Jarrett, & Trivedi, 1996) on 14 occasions, and pre-/post-CT measures of social-interpersonal functioning and negative cognitive content.

Results—The three symptom measures marked the same severity and change constructs, and we offer improved formulas for inter-measure score conversions via their common factor. Pre-post CT symptom reductions were large (d s 1.71-1.92), and nomothetic symptom curves were log-linear (larger improvements earlier and smaller improvements later in CT). Nonetheless, only 30% of individual patients showed clear log-linear changes, whereas other patients showed linear (e.g., steady decreases; 20%), one-step (e.g., a quick drop; 16%), and unclassified (34%) patterns. Log-linear, linear, and one-step patients were generally similar to one another and superior to unclassified patients post-CT in symptom levels, response and stable remission rates, social-interpersonal functioning, and cognitive content (median $d = 0.69$).

Conclusions—Reaching a low-symptom “destination” at the end of CT via any coherent “path” is more important in the short-term than which path patients take. We discuss implications for theories of change, clinical monitoring of individuals’ progress in CT, and the need to investigate long-term outcomes of patients with differing symptom change patterns.

Keywords

major depressive disorder; cognitive therapy; symptom assessment; change pattern

Acute-phase cognitive therapy (CT; Beck, Rush, Shaw, & Emery, 1979) is used to reduce symptom severity for patients with major depressive disorder (MDD), first to achieve a response (i.e., moderate or low depressive symptoms and no MDD) and then full remission (i.e., asymptomatic; Frank et al., 1991). Roughly two-thirds of patients who complete CT respond, and completers experience a large average reduction in depressive symptom severity (typically 1-3 *SD*; e.g., Craighead, Sheets, Brosse, & Ilardi, 2007; Vittengl, Clark, Kraft, & Jarrett, 2005). Although encouraging, these summary statistics can obscure differences among methods of measuring depressive symptoms, variability in the average rate of symptom change during treatment, and patient-to-patient variability in patterns of symptom change. These obscurities make it difficult to know how the results of a given CT study measuring depressive symptoms with the Beck Depression Inventory (BDI; Beck, Ward, Mendelson, Mock, & Erbaugh 1961) compared to those of another using the Inventory for Depressive Symptomatology—Self-Report (IDS-SR; Rush, Gullion, Basco, Jarrett, & Trivedi, 1996) or Hamilton Rating Scale for Depression (HRSD; Hamilton, 1960). Pre-/post-CT comparisons also do not tell us how quickly the average CT patient improves, how many patients show coherent patterns of change that deviate from the average, or how outcomes vary among patients with different change patterns.

The current research addresses these issues in a large sample of patients with recurrent MDD treated with CT and includes three measures of depressive symptoms (HRSD, BDI, IDS-SR) at 14 repeated assessments. Our analyses focus on the subset of patients ($N = 362$ of 523) who provided sufficient data for time-series analyses by completing the CT protocol in a multi-site clinical trial (Jarrett & Thase, 2010). We first aim to replicate Vittengl, Clark, Kraft, et al. (2005) in identifying symptom-change constructs and mean patterns. We then extend this and other previous research by testing individual deviations from the mean change pattern and comparing outcomes in several domains among patients following different change patterns.

The relation between the amount of treatment provided (e.g., number of psychotherapy sessions or time in treatment) and symptom level may take several shapes. Arguably the simplest shape is linear, with equal decreases in symptoms from session-to-session or week-to-week in treatment. Steady linear improvements might represent incremental learning and practice of skills taught in psychotherapy. For example, among German patients with mixed diagnoses receiving a variety of individual psychotherapies, Percevic, Lambert, and Kordy (2006) found that a linear change model described patients' mean improvements as well as more complex curvilinear models. Similarly, Barkham, Stiles, and Shapiro (1993) found that linear relations between time in treatment and problem ratings described 72% of depressed and anxious patients' responses to psychodynamic-interpersonal or cognitive-behavioral psychotherapies. Nonetheless, Barkham et al. noted substantial individual differences, including curvilinear change in about 28% of patients. Interestingly, these same approximate proportions of one-quarter curvilinear (rapid early recovery) and three-quarters linear (gradual improvement) characterized another cohort of patients' responses to antidepressant pharmacotherapy (Uher et al. 2010).

Although linear improvements characterize many patients, Howard, Kopta, Krause, and Orlinsky (1986) concluded via meta-analysis that the average dose-response curve (number of psychotherapy sessions plotted against the probability of improvement) is a decelerating, log-linear function—the first few sessions are associated with greater gains than are later sessions. The log-linear model has been replicated across several diagnoses, types of psychotherapy, and symptom measures (Lutz, Martinovich, Howard, & Leon, 2002). Quicker, earlier improvements during treatment have been attributed to “remoralization” or restoration of hope, whereas slower, later reductions in symptoms may reflect learning and practicing therapy skills (Howard, Lueger, Maling, & Martinovich, 1993). Similarly, Beck et

al. (1979) initially conceptualized early CT sessions as focusing on quicker symptom reduction and later sessions as focusing on relapse prevention. Average dose-response curves may provide means to benchmark patients' progress (e.g., is this patient making adequate progress or should the treatment be discontinued or changed?) and to recommend minimum standard lengths of treatment (e.g., Hansen, Lambert, & Forman, 2002; Lutz et al., 2002). Unfortunately, most trials in routine practice may be too short to produce clinically significant change (e.g., perhaps 13-18 sessions are needed for half of patients to improve but most patients receive fewer than 5 sessions; Hansen et al., 2002).

Applications of dose-response models have been challenged by the finding that, in naturalistic datasets where the number of psychotherapy sessions delivered is flexible, patients who receive fewer sessions versus longer courses of treatment tend to improve more quickly versus slowly, respectively (e.g., Baldwin, Berkeljon, Atkins, Olsen, & Nielsen, 2009; Barkham et al., 2006). One interpretation is that patients and/or therapists often decide to stop treatment when patients' improvements are judged "good enough." Weak correlations between symptom changes earlier versus later in treatment also challenge benchmarking patients' progress by dose-response curves. Percevic et al. (2006), for example, found that earlier improvements were weak *negative* predictors of later improvements ($r = -.22$) in sample of patients with mixed diagnoses receiving unspecified individual psychotherapies in a German hospital. Similarly, early and late changes in depressive symptom levels formed independent factors in a sample of outpatients with recurrent depression who received CT (Vittengl, Clark, Kraft, et al., 2005). Percevic et al. (2006) suggested that current symptom level, rather than a patient's recent progress or lack of progress, is most important in outcome monitoring—if current symptom levels are too high (not "good enough") then treatment should continue. The literature is unclear, however, regarding at what point it is better to continue the same treatment or to augment or switch treatments.

In contrast to linear or log-linear symptom curves, abrupt "insight" may produce a one-step drop in symptom scores (Caspar & Berger, 2007). Although the concept of insight has psychodynamic origins, it can be applied broadly to reflect that the patient has learned some concept. In CT, concepts to be learned (i.e., insights to be gained) include that thoughts, which can be changed, influence emotion (Grosse Holtforth et al., 2007). Studying patterns of symptom change that may signal patients' learning of (i.e., gaining insight into) key concepts may clarify how CT works (Elliott, 2010). For example, patients with "sudden gains" (e.g., Tang & DeRubeis, 1999) and rapid early response (e.g., Hayes, Feldman, et al., 2007) show better functioning at the end of CT for depression compared to patients without these abrupt changes. Building on past research, we implemented a broader, more flexible method that does not restrict identification of the one-step pattern to a particular time or size (e.g., rapid early response in the first 4 weeks of treatment marking at least 60% of the patient's total symptom improvement, Ilardi & Craighead, 1994; sudden gains no sooner than between the second and third CT sessions and of at least 7 BDI points, Tang & DeRubeis, 1999). Instead, our one-step method tests for a stable change between adjacent assessments at any time in treatment and is concerned with the orderliness of the step (i.e., within-patient variance accounted for) rather than a partly arbitrary size threshold. Further, our one-step method controls the Type I error rate when assessing changes in multiple intervals, which is absent from study of sudden gains (Tang & DeRubeis, 1999). Finally, research focusing on a single change pattern, such as sudden gains, may inflate its apparent benefits by ignoring other coherent change patterns and by including patients with no systematic change (e.g., non-responders) in the comparison group. In the current analyses, we compared patients with the one-step pattern to other coherent change trajectories.

The current analyses aim to replicate, integrate, and extend previous work on symptom-change patterns in patients with MDD receiving CT. We first confirmed that patients' BDI, HRSD, and IDS-SR scores improve substantially and similarly during CT, and that all mark the same symptom severity and change constructs (i.e., cross-sectionally and longitudinally). This confirmation is necessary so that conclusions about symptom change trajectories apply to depressive symptoms generally rather than only to a particular measure (e.g., the BDI vs. ID-SSR) or method (e.g., patient vs. clinician reports). We then tested individual differences in symptom-change patterns, estimating the frequency of linear, log-linear, and one-step symptom-change patterns. Next, we explored differences among patients with—and without—one of these coherent change patterns pre- and post-CT, considering levels of depressive symptoms, depressive cognitive content, and social-interpersonal functioning. We did not make directional hypotheses because both theory and prior empirical tests about the superiority of competing change patterns in CT are sparse. Finally, among patients not showing clear linear, log-linear, and one-step patterns, we explored additional possible symptom-change patterns.

Method

The current analyses use data from the acute phase of a clinical trial comparing the relapse/recurrence rates of acute-phase CT responders randomized to continuation-phase CT, fluoxetine, or pill placebo (Jarrett & Thase, 2010) called the Continuation Phase Cognitive Therapy Relapse Prevention (C-CT-RP) Trial, which is registered at ClinicalTrials.gov (NCT00118404, NCT00183664, and NCT00218764). This research was approved by Institutional Review Boards at both data collection sites, The University of Texas Southwestern Medical Center at Dallas and University of Pittsburgh Medical Center. Participants provided written informed consent for evaluation and treatment. Participants were withdrawn from psychotropic medications before entering the study and were not prescribed antidepressant medications in the acute-phase CT protocol. Below we summarize the methods relevant to the current analyses and refer readers to Jarrett and Thase (2010) for additional detail, including description of the continuation and follow-up phases not examined in this report.

Participants

Participants were self- or practitioner-referred outpatients who (a) met *DSM-IV* criteria for recurrent MDD (American Psychiatric Association, 2000), (b) had previously remitted between depressive episodes, had at least one prior depressive episode with complete inter-episode recovery, or had antecedent dysthymic disorder, and (c) scored ≥ 14 on the 17-item HRSD.¹ Excluded participants: a) had severe or poorly controlled concurrent medical disorders that could cause depression, b) had psychotic or organic mental disorders, bipolar disorder, active substance dependence, or primary obsessive-compulsive or eating disorders, c) could not complete questionnaires in English, d) represented an active suicide risk, e) were <18 or >70 years old, f) had not responded previously to ≥ 8 weeks of CT or 6 weeks of fluoxetine, or g) were pregnant or planned to become pregnant during the first 11 months after intake. Psychiatric diagnoses were made via the Structured Clinical Interview for *DSM-IV* (First, Spitzer, Gibbon, & Williams, 1996). Similar to the full sample of consenting patients ($N = 523$; Jarrett & Thase, 2010), the 362 patients with sufficient data (completed 39 of 42 target symptom assessments) for analysis in this report had age $M = 43.4$ years ($SD = 12.2$); 69.1 % were women; completed $M = 15.5$ ($SD = 2.8$) years of education; 1.9% were

¹Two patients erroneously entered CT with HRSD = 13 at one of two diagnostic visits. During CT, one of these patients responded and one dropped out. As recommended by the Data Safety and Monitoring Board (DSMB), the two patients are analyzed here as they were treated during data collection.

Asian, 6.9% black, 4.7% Hispanic, 85.3% white, and 1.1% other ethnicities. Patients' MDD age of onset was $M = 21.2$ years ($SD = 10.6$) and their current major depressive episode had lasted $M = 26.7$ months ($SD = 47.7$).

Acute-Phase CT

Sixteen therapists who demonstrated competence by achieving and maintaining Cognitive Therapy Scale (CTS; Young & Beck, 1980) scores ≥ 40 provided CT. Therapists submitted videotaped sessions for review and received weekly group supervision. Supervisors and other therapists rated videotaped sessions on the CTS to provide therapists with feedback on strengths and weaknesses. The CT protocol included 16 or 20 sessions spread over 12 (maximum 14) weeks. Patients received 2 sessions weekly for 4 weeks. After 4 weeks, patients who achieved $\geq 40\%$ reduction in the HRSD score compared to diagnostic follow-up received 1 session weekly for 8 weeks, whereas patients with $< 40\%$ reduction in HRSD received 2 sessions weekly for 4 weeks and then 1 session weekly for 4 weeks to maximize their chances of response and eligibility for participation in later phases of the trial.² Patients with data analyzed in this report completed the protocol by attending 14 (of 16) or 18 (of 20) CT sessions.

Measures

Depressive symptom severity—Patients completed the 21-item BDI and 30-item IDS-SR, and clinicians administered the 17-item HRSD at the diagnostic evaluation, weekly during CT, within 7 days of the last CT session, and any time a patient exited the protocol. Higher scores indicate more severe depressive symptoms, and the measures' validity for assessing depressive symptoms is well established (e.g., Vittengl, Clark, Kraft, et al., 2005). Alpha internal consistency reliability was acceptable for the BDI (median = .91), IDS-SR (median = .90), and HRSD (median = .82). Based on a multilevel analysis of 28 patients rated by 4-14 clinicians each, the HRSD demonstrated interrater reliability of ICC = .91 in the current study.

Major depressive episode (MDE)—The Current MDD section of the SCID-I for *DSM-IV* was administered at the diagnostic evaluation, within 7 days of the last CT session, and any time a patient exited the protocol. Inter-rater reliability for MDE was moderate in the current dataset (ICC = .61 in a multilevel analysis of 39 patients rated by 4 to 21 clinicians each).

Psychosocial functioning—Patients rated their psychosocial functioning on the Social Adjustment Scale—Self-Report (SAS-SR; Weissman & Bothwell, 1976) and Inventory of Interpersonal Problems (IIP; Horowitz, Rosenberg, Baer, Ureño, & Villaseñor, 1988). On the 56-item SAS-SR, patients rated their dysfunction in relevant domains (e.g., work, leisure, parental, marital, social) on a scale from 1-5. On the 127-item IIP, patients rated from 0-4 the extent to which behaviors, thoughts, and feelings have been problematic in their significant relationships. Higher mean scores on the SAS-SR and IIP indicate poorer functioning. In support of the SAS-SR's validity, Weissman, Olfson, Gameroff, Feder, and Fuentes (2001) reported means of 2.5 for depressed and 1.7 for non-depressed samples. Validity data for the IIP include sensitivity to patients' improvement in psychotherapy (Horowitz et al., 1988). At the intake and post-CT assessments used in current study, the SAS-SR (.73, .81) and IIP (.96, .98) total scores showed acceptable alpha internal consistency reliability, respectively.

²Four (1 early and 3 late) responders were misclassified as late and early responders, respectively. As recommended by the DSMB, they are analyzed here as they were treated during data collection.

Cognitive content—Patients completed the Dysfunctional Attitudes Scale (DAS; Form A; Weissman, 1979) and the Beck Hopelessness Scale (BHS; Beck, Weissman, Lester, & Trexler, 1974). On the 40-item DAS, patients rated statements about self-concept, happiness, perfectionism, and thoughts and feelings on a 7-point scale of agreement. On the 20-item BHS, patient rated their expectancies about the future on a true/false scale. Higher total scores on the DAS and BHS reflect greater negative cognitive content. The validity of the DAS is supported by concurrent and predictive correlations with depressive symptoms (Otto et al., 2007), and higher scores on the BHS predict greater suicidality (Brown, Beck, Steer, & Grisham, 2000). At the intake and post-CT assessments used in current study, the DAS (.93, .94) and BHS (.89, .92) total scores showed acceptable alpha internal consistency reliability, respectively.

Estimation of Missing Symptom Data

We chose to analyze the subsample of patients with mostly-complete depressive symptom data to balance competing goals of maximizing the generalizability of results (by including more patients) with effective modeling of symptom change trajectories (by including patients with symptom data at more time points). At one extreme, an intent-to-treat analysis would include all patients, including patients who were offered treatment at intake and attended few, or even no, CT sessions. Thus, some patients in an intent-to-treat analysis would have data at only 1 or 2 time points making estimation of the trajectories difficult or impossible (e.g., linear and non-linear trajectories cannot be differentiated with only 2 time points). The other extreme would be to include only patients with complete data in the analysis, but we rejected this strategy because it excludes patients who are missing even one symptom score. Instead, we examined a frequency distribution of missing data and found that we could increase the sample size from 266 to 362 patients with complete data by estimating 3 depressive symptom scores (of a possible 42) per patient. We estimated these values by averaging 5 imputations generated via the Markov Chain Monte Carlo method (Schafer, 1997), an iterative technique that uses observed means and covariances. Although we present only analyses including estimated data, analyses without estimated data supported substantively equivalent conclusions.

Results

Symptom Severity and Change during CT

To allow robust conclusions about changes in the construct of depressive symptoms, we first established that the three measures (HRSD, BDI, and IDS-SR) marked the same construct. We tested interrelations of the HRSD, BDI, IDS-SR over 14 assessments (diagnostic *Intake*, *Weeks 1-12* of CT, and a follow-up completed within 1 week *Post-CT*) with principal-axis factor analyses (methods detailed by Vittengl, Clark, Kraft et al., 2005). We analyzed zero-order correlations of symptom severity scores and partial correlations controlling intake scores to examine changes in symptom severity. Replicating Vittengl, Clark, Kraft, et al. (2005), both analyses supported two- and three-factor varimax-rotated solutions (proportion of common variance accounted for by factors 1-3 = 65.6%, 13.5 %, 5.0%, and 63.5 %, 13.2%, 5.8 %, in zero-order and partial correlation analyses, respectively). We present the rotated two-factor solutions for their greater simplicity and clarity (Table 1) because time was the critical organization variable in the analyses.³ The two-factor solutions included “early” (*Intake* through roughly *Week 6*) and “late” (roughly *Week 7* through *Post*) factors on which all three measures loaded strongly. These analyses confirmed that the HRSD, BDI, IDS-SR mark the same depressive symptom construct (e.g., median convergent $r = .81$),

³Loadings for the three-factor solutions, as well as zero-order and partial correlation matrices, are available from the first author.

which has decreasing retest reliability at longer intervals (e.g., median $r = .81, .58, .35$ at lags of 1, 5, and 10 assessments, respectively).

Mean symptom change was also highly convergent among the BDI, HRSD, and IDRS. The measures showed quicker symptom change earlier in treatment (see Figure 1), and the effect sizes for reduction in symptom severity from *Intake* to *Post* were large and similar for the HRSD ($d = 1.71$; $CI_{95\%} 1.54-1.87$), BDI ($d = 1.73$; $CI_{95\%} 1.57-1.89$), and IDS-SR ($d = 1.92$; $CI_{95\%} 1.75-2.09$). Among these patients completing CT, 74% ($CI_{95\%} 69-79\%$) were judged to have clinically significant improvements by meeting criteria for response at exit (no MDE and HRSD ≤ 12), and 12% ($CI_{95\%} 9-16\%$) showed stable remission (no MDE and the last seven scores on the HRSD ≤ 6), following outcome definitions designed to identify patients at higher and lower risk for relapse/recurrence, respectively (Jarrett & Thase, 2010).

Conversion among the HRSD, BDI, and IDS-SR through Their Common Factor

Because the HRSD, BDI, and IDS-SR mark the same depressive symptom construct and change similarly during CT, clinicians and researchers using one of the measures can estimate scores on the others. We replicated Vittengl, Clark, Kraft, et al.'s (2005) procedure to compute concurrent inter-scale conversions via the scales' common factor. Specifically, we pooled data across assessments (362 patients \times 14 assessments = 5068 observations each for the HRSD, BDI, and IDS-SR) and conducted a principal axis factor analysis. Because time was controlled via the patient \times assessment data structure, unlike the multi-factor solutions above for the between-patient data, a one-factor solution accounted for 100% of the common variance and all three measures loaded highly on it (HRSD = .88; BDI = .93; IDS-SR = .95). We predicted each raw-score depression scale from the linear factor score in a linear regression analysis. The resulting regression constants and coefficients allow conversion of the depressive-symptom measures through their common factor (see Table 2). Within the score ranges observed in the current dataset (HRSD 0-37; BDI 0-59; IDS-SR 0-74), the current factor-analytic formulas agreed very strongly with Vittengl, Clark, Kraft, et al.'s (2005) formulas (median ICC = .99, range .98-1.00, among pairwise scale conversions).

Nomothetic Pattern of Change during CT

Rather than duplicate analyses with highly convergent and interchangeable BDI, HRSD, and IDS-SR symptom measures, we analyzed nomothetic and idiographic change patterns using a composite index. In particular, we standardized and averaged the measures in reference to their distributions at *Intake*. We scaled the symptom-severity composite as T-scores ($M = 50$, $SD = 10$) at *Intake*, with smaller values reflecting lower depressive symptoms.

We examined mean scores by assessment time point to understand nomothetic symptom levels. We tested linear (i.e., a consistent rate of change) and log-linear (i.e., a decelerating rate of change) models because of their parsimony and theoretical relevance (e.g., Howard et al., 1986; Percevic et al., 2006). By visual inspection (see Figure 1) confirmed with multilevel models (see Table 3), depressive symptoms followed decelerating curves with greater changes earlier, followed by smaller changes later in time. The multilevel models included both fixed effects (capturing nomothetic patterns) and random effects (capturing individual patients' deviations from the nomothetic patterns). The significant linear fixed-effect slope shown indicated that, on average, symptoms decreased 2.3 points per assessment time point. However, the log-linear model fit the data better (as shown by a lower Bayes Information Criterion [BIC] value), and the log-linear fixed-effect slope indicated that the average gains were larger earlier—and smaller later—in treatment. For example, the mean decrease in symptoms between treatment *Weeks 1* and *2* (6.4 points) was several times larger than between *Weeks 11* and *12* (0.7 points).

Idiographic Patterns of Change during CT

Although nomothetic depressive symptom means followed decelerating curves, multilevel models also revealed significant variability among individual patients. The *assessment* and *log(assessment)* random effects (see Table 3) indicated that nomothetic linear and curvilinear slopes fit patients' symptom changes to varying degrees. Although many systematic patterns of change are conceivable, we examined individual differences by testing three simple patterns that capture major ideas in psychotherapy process theories.

We fit a series of regressions to each patient's individual series of symptom composite scores.⁴ For each patient, we regressed symptom scores at the 14 assessments on linear (*assessment* number 1-14) and log-linear (*log assessment*) functions. In addition, we created a "one-step" model that captured the extent to which patients have a clear break in symptom score levels at one point in their series of assessments (e.g., due to an insight). We did this by creating 13 dummy variables representing all possible breaks in the series of 14 assessments (the first assessment represented by a 1 and the following thirteen assessments with 0s; the first two assessments represented by 1s and the following twelve assessments symptom 0s, etc.) and selecting the best predictor of symptom scores among these 13 dummy variables. For example, a patient with a series of symptom scores 51, 50, 52, 49, 50, 32, 33, 31, 32, 32, 33, 31, 30, 29 would best fit the dummy variable with values 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0 indicating a step down in symptoms after the fifth assessment.

To investigate model fit for each patient, we converted patients' raw R^2 values to percentile ranks of R^2 values derived from a simulation.⁵ In the simulation, we sampled 10,000 cases from a population with the following properties: (a) 14 repeated assessments of symptoms; (b) correlations among the repeated assessments equal to the correlation matrix for actual patients (e.g., higher correlations among closer assessments); (c) symptom M and SD at all assessments equal to the actual patients' values at *Intake* ($M = 50$, $SD = 10$). The third property of "no change in population means" allowed us to estimate the probability p that random fluctuations in symptoms would demonstrate linear, log-linear, and one-step change patterns to varying degrees. We fit linear, log-linear, and one-step models to each of the 10,000 simulated cases and retained the R^2 values from the regressions. We concluded that actual patients followed an "unclassified" pattern of change when none of their R^2 values reached the 95th percentile of linear, log-linear, or one-step R^2 values in the simulation. We categorized the remaining patients with R^2 values at or above the 95th percentile ($p < .05$) as following the pattern (linear, log-linear, or one-step) with the highest percentile fit.

Table 4 shows high R^2 values for actual patients with linear, log-linear, and one-step patterns. Not surprisingly, given the average symptom curve shown in Figure 1, the log-linear pattern was common for individual patients. Nonetheless, linear and one-step patterns

⁴Growth mixture modeling can address broadly similar goals. Growth mixture modeling is an exploratory technique that seeks to identify homogeneous groups of people with distinct growth trajectories (e.g., groups of patients with different symptom change patterns; Uher et al., 2010). As an exploratory technique, growth mixture modeling better generates than tests hypotheses. Patients grouped by growth mixture modeling are more similar in trajectory to one another than they are to patients in different groups (e.g., Jung & Wickrama, 2008). However, group membership and group-level fit statistics do not reveal whether individual patients strongly fit a hypothesized trajectory (e.g., linear or log-linear symptom reduction). Consequently, we first employed idiographic regression analyses to test individual patients' trajectory fits. For remaining patients without strong fits to hypothesized trajectories, we then used growth mixture modeling to explore additional possible change patterns in the population.

⁵It would be more convenient but less accurate to use p and R^2 values derived from the regressions directly to measure how well each pattern fit. However, the one-step model may yield inflated R^2 values compared to the linear and log-linear models because the one-step model retains only the best-fitting (among 13 possible) breaks in the data series. Further, dependencies (e.g., autocorrelation) in the series of symptom scores may bias p in complex ways in all three models (e.g., Kenny & Judd, 1986). Consequently, when these numbers are taken directly from individual patients' regression models, R^2 may not provide a valid means of comparing fits between the three competing models, and p may not accurately distinguish models that fit patients' data series better than expected by chance. Our Monte Carlo simulation was designed to support inferences about R^2 when traditional regression assumptions, especially independence of residuals, are not met (Fan, Felsovalyi, Sivo, & Keenan, 2002), as well as comparisons among competing models.

were not rare. Among patients with the one-step change pattern, breaks ranged from between *Weeks 1-2* to between *Weeks 10-11* (median between *Weeks 3-4*) of CT. It is important to note that linear, log-linear, and one-step models may strongly fit both small and large overall decreases (and even increases) in symptom scores, as long as the variability in the scores conforms to the modeled pattern. However, the negative slopes shown in Table 4 indicate that all patients with linear, log-linear, and one-step patterns improved during CT. For example, the mean one-step coefficient indicates that average symptom levels were nearly 3 *SD* (28 T-score points) lower after the step. Figure 2 shows symptom means for patients following linear, log-linear, and one-step changes.

Predictors and Outcomes of Idiographic Change Patterns

We compared demographic, clinical, social functioning, and cognitive content variables for patients showing linear, log-linear, one-step, and unclassified patterns of change in depressive symptom composite scores (see Table 5). Based on one-way ANOVAs for continuous variables and exact tests for categorical variables, the four groups of patients showed no differences at *Intake* in demographic and illness characteristics, symptom severity, social-interpersonal functioning, or cognitive content (p s > .05).

In contrast, at the *Post* assessment completed within a week of the end of CT, moderate to large differences were evident (see Table 5). The linear, log-linear, and one-step patient groups, respectively, all showed better outcomes compared to the unclassified group by having lower depressive symptoms (d s 1.18 [0.87-1.50], 1.12 [0.84-1.40], and 0.93 [0.61-1.26], [CI_{95%}]), greater proportions of patients responding to treatment (d s 1.13 [0.82-1.44], 0.99 [0.72-1.26], and 0.96 [0.63-1.28]), greater proportions of patients with stable remission of depression (d s 0.42 [0.12-0.71], 0.70 [0.43-0.96], and 0.96 [0.63-1.29]), better social adjustment (d s 0.79 [0.48-1.11], 0.68 [0.41-0.96], and 0.69 [0.35-1.02]), fewer interpersonal problems (d s 0.51 [0.20-0.82], 0.51 [0.24-0.78], and 0.40 [0.07-0.73]), fewer dysfunctional attitudes (d s 0.66 [0.35-0.98], 0.35 [0.08-0.63], and 0.63 [0.30-0.96]), and less hopelessness (d s 0.97 [0.65-1.28], 0.62 [0.35-0.90], and 0.67 [0.33-1.00]). The linear, log-linear, and one-step groups did not differ from one another in depressive symptoms, response proportions, social-interpersonal functioning, and dysfunctional attitudes. However, the linear group reported less hopelessness than the log-linear group ($d = 0.36$ [0.05-0.67]). In contrast, the linear group showed stable remission less frequently than the log-linear ($d = 0.34$ [0.04-0.64]) and one-step ($d = 0.48$ [0.13-0.83]) groups because the linear group patients did not reach their lowest symptom levels as quickly.

Exploration of “Unclassified” Idiographic Change Patterns

We tested linear, log-linear, and one-step patterns based on a priori theory, but about one-third of patients did not demonstrate any of these patterns clearly. To explore possible additional systematic change patterns, we submitted unclassified patients' symptom scores to growth mixture model analysis using PROC TRAJ software (Jones & Nagin, 2007). We evaluated models with 1-5 subgroups: intercept-only (no change), linear (straight line trajectory), quadratic (one curve or bend in the trajectory), and cubic (two curves or bends) functions. We compared models using BIC, favoring models with lower values (i.e., better fits; Jones & Nagin, 2007).

A four-group model with intercept-only, linear, quadratic, and cubic groups fit the data comparatively well. Figure 3 shows the groups' average symptom scores and model coefficients. Group 1 (cubic change trajectory) showed moderate improvements, especially early in treatment, and 70% of these patients were responders to CT. Group 2 (quadratic change trajectory) patients showed modest improvements during the first half of CT, but made little progress later, and only 45% were responders. Group 3 (linear change trajectory)

patients showed small improvements throughout CT and only 9% were responders. Finally, Group 4 (no change trajectory) patients had high symptom levels at *Intake* ($M = 66$), showed little improvement, and none (0%) were responders. Although the mean scores shown in Figure 3 for Groups 1 and 2 appear roughly log-linear, and Group 3 is roughly linear, it is important to note that individually these patients had too much quasi-random variation in symptom levels from assessment to assessment to demonstrate coherent linear, log-linear, or one-step patterns in primary analyses. For example, the mean R^2 fit for patients in the linear group identified in the growth mixture model (.32) was much poorer than for patients identified as showing linear change in the primary analyses (.87). Consequently, the trajectories identified via the growth mixture models likely are less robust than trajectories identified in the primary analyses.

Discussion

The current analyses clarified both nomothetic and idiographic patterns of change in depressive symptoms during CT for MDD. Replicating past research (Vittengl, Clark, Kraft, et al., 2005), we showed in an independent sample that three common depression symptom measures (clinician-rated HRSD and patient-rated BDI and IDS-SR) reflect the same symptom level and change constructs throughout CT and so can be aggregated. We offer formulas for clinicians and researchers to convert among the HRSD, BDI, and IDS-SR and, although highly convergent, we recommend them over formulas published earlier (Vittengl, Clark, Kraft, et al., 2005) because the current sample is nearly three times larger. For example, a review suggests that depressed patients' average post-CT score on the BDI is about 9 (Craighead et al., 2007), and clinicians using the current formulas could estimate that their patients should score about 15 on the freely-available IDS-SR.

Patients' mean depressive symptoms followed a decelerating, log-linear curve with larger improvements earlier, and smaller improvements later, in CT. Although decelerating nomothetic change curves are common (e.g., Lutz et al., 2002; Howard et al., 1986), most idiographic changes did not fit this pattern clearly. Instead, many individual patients showed coherent linear and one-step improvements during CT. Post-CT outcomes for patients with any of these three patterns were superior to those for patients without coherent change, in terms of symptom severity, odds of treatment response and stable remission, cognitive content, and social-interpersonal functioning. Because patients with linear, log-linear, and one-step symptom change trajectories showed similar short-term outcomes, the possible mechanisms generating the trajectories should be valued equally. For example, the data would not support an argument that one-step patterns reflecting "insights" are superior to linear improvements that reflect "piecemeal learning" or log-linear change reflecting "restoration of hope and consolidation of gains" in short-term outcomes.

The current study extends past research that compared patients with one coherent change pattern to all other patients (e.g., sudden gains, a special case of the one-step pattern; Tang & DeRubeis, 1999; Vittengl, Clark, & Jarrett, 2005), including patients without coherent change, by differentiating several coherent change patterns. Future research should compare the long-term outcomes of patients with these three change patterns. That is, although the short-term outcomes were similar, the superiority of one pattern may be evident later in lower relapse rates or improved psychosocial functioning. Parallel to findings of increased risk for relapse with unstable (vs. stable) response to CT (Jarrett et al., 2001; Jarrett et al., in preparation), responders without coherent linear, log-linear, or one-step change patterns may be more susceptible to relapse and recurrence than patients with coherent change patterns.

The processes producing different symptom change trajectories in CT are unknown and possible topics for future research to clarify mechanisms of change (Hayes, Laurenceau,

Feldman, Strauss, & Cardaciotto, 2007; Kazdin, 2007). For example, research might test interactions of patient characteristics (e.g., level of reward-seeking behavior, degree of depressive content in automatic thoughts and schema) with the early (e.g., behavioral activation), middle (e.g., cognitive assessment and restructuring of negative automatic thoughts), and later (e.g., schema work) components of CT (Beck et al., 1979). Perhaps matches of individual patients' weaknesses (or strengths) with therapists' delivery of therapy components produce larger reductions in symptoms at different stages in treatment. At the same time, future research might also track external psychosocial functioning associated with changes in depressive symptoms (e.g., Dunn et al., 2012) and consider rapid early response phenomena that may relate to common factors such as therapeutic alliance (e.g., Ilardi & Craighead, 1994). Finally, research might address whether subtypes of MDD (e.g., chronic vs. recurrent) or comorbid disorders (e.g., anxiety, personality) relate to different change trajectories and mechanisms. For example, it is unknown whether the current sample of patients with recurrent MDD evidenced more one-step and fewer linear change trajectories than would be found in among patients with chronic MDD.

The current results present challenges to clinical tracking and prognostication. Because the three identified coherent change patterns predicted similar short-term outcomes, it may be difficult to gauge whether individual patients are "on track," especially early in CT (cf. Percevic et al., 2006). Although the odds of response decrease the longer patients maintain high symptom levels during CT (Thase et al., in preparation), this may be due, at least in part, to use of a time-limited protocol, rather than indicating highly stable causes of non-response. For example, early and late depressive symptoms were largely independent in the current analyses; that is, depressive symptoms were highly correlated concurrently (median $r = .81$) and across short intervals (lag 1 median retest $r = .81$) but much less stable over longer periods (lag 10 median retest $r = .35$). Thus, lower response rates in patients who did not show symptom change by mid-treatment may simply indicate that as CT progresses there is less opportunity to improve within a limited number of remaining CT sessions. In addition, no tested intake variables, including demographics, prior course of MDD, early symptom levels, cognitive content, and social-interpersonal functioning predicted CT symptom change patterns. That is, our data did not reveal at intake which path to CT response, if any, patients would follow.

The one-third of patients who did not show coherent linear, log-linear, or one-step change patterns formed distinguishable groups in exploratory growth mixture models. Analogous to dividing residual Not Otherwise Specified (NOS) diagnostic categories into positive diagnoses (e.g., differentiating Binge Eating Disorder from Eating Disorder NOS; American Psychiatric Association, 2000), the change pattern groups identified post-hoc may have some external validity. For example, the largest post-hoc group followed a cubic symptom curve, and most of these patients met response criteria for CT. Patients in the other three post-hoc groups showed less (linear, quadratic) or no (intercept-only) improvement, on average, and fewer met criteria for response. As follow-up data become available, responders who continued in the current clinical trial could be compared to understand the importance, if any, of these post hoc groups.

Characteristics of the current sample, treatment protocol, and analyses may limit our results' generalizability. First, patients had carefully diagnosed MDD treated by closely supervised cognitive therapists trained to competence criteria, unlike routine clinical practice. Although the a priori change patterns studied here are not unique to CT (e.g., some patients taking antidepressant medication also show linear and log-linear change, Uher et al., 2010; and some patients taking pill placebo show one-step change, Vittengl, Clark, & Jarrett, 2005), the short- and long-term sequelae of these change patterns may differ by diagnosis and treatment. Second, differentiating symptom trajectories required patients with mostly

complete data over the CT trial. It is worth recalling that patients who dropped out are not represented in this report. Consequently, the symptom change trajectories we identified do not predict attrition from CT, and comparisons among trajectories may not generalize to patients who do not complete CT. Third, we modeled linear, log-linear, and one-step patterns separately but some patients may show more complex but coherent changes (e.g., rapid early response followed by a transient spike in symptoms; Grosse Holtforth et al., 2012; Hayes, Feldman, et al., 2007). Future research could examine more complex patterns when justified based on theory or clinical observation. Fourth, our symptom trajectory analyses used data from 14 weekly assessments, which meets or exceeds data collection in most clinical trials, but may not provide adequate statistical power to detect some patients' coherent change patterns (e.g., may need 50 assessments; Barlow, Nock, & Hersen, 2009) and makes regression analyses more sensitive to violations of assumptions (e.g., inflate or deflate R^2 due to heteroscedasticity; Cohen, Cohen, West, & Aiken, 2003). Finally, in our dose-response analyses, we modeled changes in continuous depressive symptom scores taken weekly. Because researchers have found similar decelerating response curves by modeling dose as number of psychotherapy sessions and/or response as odds of dichotomous improvement (e.g., Howard et al., 1986), we believe that our nomothetic pattern of results is robust. But it would be difficult (if not impossible) to model linear, log-linear, and one-step idiographic change patterns using a dichotomous weekly measure.

When treatments are effective, symptom levels decrease over time. Because CT is an effective treatment for depression (e.g., Craighead et al., 2007), nomothetic and coherent idiographic changes during treatment usually reflect decreases in symptoms. Which coherent symptom-decrease path patients take may be less important than taking a coherent vs. quasi-random path. Assuming that CT's active processes are reflected in weekly symptom scores, the current results suggest two non-exclusive possibilities: Several therapeutic processes in CT produce different change patterns (e.g., common factors such as placebo-expectancy effects or the therapeutic alliance, CT-skills based responses), or extra-therapeutic conditions (e.g., external life events, patient characteristics not tested in the current analyses) moderate weekly changes in symptoms produced by CT. Addressing these possibilities in future research may clarify how CT produces its well-established benefits.

Acknowledgments

This report was supported by Grants Number K24 MH001571, R01 MH58397, R01 MH69619 (to Robin B. Jarrett, Ph.D.) and R01 MH58356 and R01 MH69618 (to Michael E. Thase, M.D.) from the National Institute of Mental Health (NIMH). In addition, this report was supported by a faculty research grant to Jeffrey R. Vittengl, Ph.D., from the School of Social and Cultural Studies, Truman State University. We appreciate the support of our NIMH Program Officer, Jane Pearson, Ph.D. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIMH or the National Institutes of Health. We also appreciate the careful review by members of the trial's Data Safety and Monitoring Board. We are indebted to our research teams and our colleagues at The University of Texas Southwestern Medical Center at Dallas, the University of Pittsburgh (where Dr. Thase was located during patient accrual), and the University of Pennsylvania (Dr. Thase's current affiliation). We appreciate the assistance of Joanne Sanders, M.S., and Abu Minhajuddin, Ph.D., in preparing this manuscript.

Drs. Vittengl and Clark have no interests to disclose regarding this research. Dr. Thase has no conflicts of interest pertaining to this paper, although he does report the following relationships with companies that develop treatment for depression or provide education pertaining to those treatments: Dr. Thase has provided scientific consultation to Alkermes, Astra-Zeneca, Bristol-Myers Squibb Company, Dey Pharma, L.P., Eli Lilly & Company, Forest Pharmaceuticals, Inc., Gerson Lehman Group, GlaxoSmithKline, Guidepoint Global, H. Lundbeck A/S, MedAvante, Inc., Merck and Co. Inc., Neuronetics, Inc., Novartis, Otsuka, Ortho-McNeil Pharmaceuticals, PamLab, L.L.C., Pfizer (formerly Wyeth-Ayerst Laboratories), Schering-Plough (formerly Organon, Inc.), Shire US Inc., Sunovion Pharmaceuticals, Inc., Takeda (Lundbeck), and Transcept Pharmaceuticals. Dr. Thase receives grant funding from the Agency for Healthcare Research and Quality, Eli Lilly & Company, GlaxoSmithKline (ended 7/10), National Institute of Mental Health, Otsuka Pharmaceuticals, and Sepracor, Inc. He has equity holdings in MedAvante, Inc. and receives royalty income from American Psychiatric Foundation, Inc., Guilford Publications, Herald House, Oxford University Press, and W.W. Norton & Company. His wife is employed as the

Group Scientific Director for (Embryon – formerly Advogent; which does business with BMS and Pfizer/Wyeth). Dr. Jarrett's medical center receives the fees from the cognitive therapy she provides to patients. Dr. Jarrett is a paid consultant to the NIMH.

References

- American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders DSM-IV-TR (Text Revision). Author; Washington, DC: 2000.
- Baldwin SA, Berkeljon A, Atkins DC, Olsen JA, Nielsen SL. Rates of change in naturalistic psychotherapy: Contrasting dose-effect and good-enough level models of change. *Journal of Consulting and Clinical Psychology*. 2009; 77:203–211. doi:10.1037/a0015235. [PubMed: 19309180]
- Barkham M, Connell J, Stiles WB, Miles JV, Margison F, Evans C, Mellor-Clark J. Dose-effect relations and responsive regulation of treatment duration: The good enough level. *Journal of Consulting and Clinical Psychology*. 2006; 74:160–167. doi:10.1037/0022-006X.74.1.160. [PubMed: 16551153]
- Barkham M, Stiles WB, Shapiro DA. The shape of change in psychotherapy: Longitudinal assessment of personal problems. *Journal of Consulting and Clinical Psychology*. 1993; 61:667–677. doi: 10.1037/0022-006X.61.4.667. [PubMed: 8370863]
- Barlow, DH.; Nock, MK.; Hersen, M. Single case experimental designs. 3rd Ed.. Pearson; New York: 2009.
- Beck, AT.; Rush, AJ.; Shaw, BF.; Emery, G. Cognitive therapy of depression. Guilford Press; New York: 1979.
- Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J. An inventory for measuring depression. *Archives of General Psychiatry*. 1961; 4:561–571. [PubMed: 13688369]
- 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. doi:10.1037/h0037562. [PubMed: 4436473]
- Brown GK, Beck AT, Steer RA, Grisham JR. Risk factors for suicide in psychiatric outpatients: A 20-year prospective study. *Journal of Consulting and Clinical Psychology*. 2000; 68:371–377. doi: 10.1037/0022-006X.68.3.371. [PubMed: 10883553]
- Caspar, F.; Berger, T. Insight and Cognitive Psychology.. In: Castonguay, LG.; Hill, C.; Castonguay, LG.; Hill, C., editors. *Insight in psychotherapy*. American Psychological Association; Washington, DC US: 2007. p. 375-399. doi:10.1037/11532-018
- Cohen, J.; Cohen, P.; West, SG.; Aiken, LS. Applied multiple regression/correlation analysis for the behavioral sciences. 3rd ed.. Lawrence Erlbaum; Mahwah, NJ: 2003.
- Craighead, WE.; Sheets, ES.; Brosse, AL.; Ilardi, SS. Psychosocial treatments for major depressive disorder.. In: Nathan, PE.; Gorman, JM., editors. *A guide to treatments that work*. 3rd Ed.. Oxford; New York: 2007. p. 289-307.
- Dunn TW, Vittengl JR, Clark LA, Carmody T, Thase ME, Jarrett RB. Change in psychosocial functioning and depressive symptoms during acute-phase cognitive therapy for depression. *Psychological Medicine*. 2012; 42:317–326. doi: 10.1017/S0033291711001279.
- Elliott R. Psychotherapy change process research: Realizing the promise. *Psychotherapy Research*. 2010; 20:123–135. doi:10.1080/10503300903470743. [PubMed: 20099202]
- Fan, X.; Felsovalyi, A.; Sivo, SA.; Keenan, SC. SAS for Monte Carlo studies: A guide for quantitative researchers. SAS Institute; Cary, NC: 2002.
- First, MB.; Spitzer, RL.; Gibbon, M.; Williams, JB. Structured Clinical Interview for DSM-IV Axis I Disorders-Patient Edition (SCID-I/P, Version 2.0). New York State Psychiatric Institute, Biometrics Research Department; New York, NY: 1996.
- Frank E, Prien RF, Jarrett RB, Keller MB, Kupfer DJ, Lavori PW, Weissman MM. Conceptualization and rationale for consensus definitions of terms in major depressive disorder: Remission, recovery, relapse, and recurrence. *Archives of General Psychiatry*. 1991; 48:851–855. [PubMed: 1929776]
- Grosse Holtforth, M.; Castonguay, LG.; Boswell, JF.; Wilson, LA.; Kakouros, AA.; Borkovec, TD. Insight in Cognitive-Behavioral Therapy.. In: Castonguay, LG.; Hill, C.; Castonguay, LG.; Hill,

- C., editors. *Insight in psychotherapy*. American Psychological Association; Washington, DC US: 2007. p. 57-80. doi:10.1037/11532-003
- Grosse Holtforth M, Hayes AM, Sutter M, Wilm K, Schmied E, Laurenceau J-P, Caspar F. Fostering cognitive-emotional processing in the treatment of depression: A preliminary investigation in Exposure-Based Cognitive Therapy (EBCT). *Psychotherapy and Psychosomatics*. 2012; 81:259–260. doi: 10.1159/000336813. [PubMed: 22678232]
- Hamilton M. A rating scale for depression. *Journal of Neurology, Neurosurgery, and Psychiatry*. 1960; 23:56–61.
- Hansen NB, Lambert MJ, Forman EM. The psychotherapy dose-response effect and its implications for treatment delivery services. *Clinical Psychology: Science and Practice*. 2002; 9:329–343. doi: 10.1093/clipsy/9.3.329.
- Hayes AM, Feldman GC, Beevers CG, Laurenceau J, Cardaciotto L, Lewis-Smith J. Discontinuities and cognitive changes in an exposure-based cognitive therapy for depression. *Journal of Consulting and Clinical Psychology*. 2007; 75:409–421. doi:10.1037/0022-006X.75.3.409. [PubMed: 17563158]
- Hayes AM, Laurenceau J, Feldman G, Strauss JL, Cardaciotto L. Change is not always linear: The study of nonlinear and discontinuous patterns of change in psychotherapy. *Clinical Psychology Review*. 2007; 27:715–723. doi:10.1016/j.cpr.2007.01.008. [PubMed: 17316941]
- Horowitz LM, Rosenberg SE, Baer BA, Ureño G, Villaseñor VS. Inventory of interpersonal problems: Psychometric properties and clinical applications. *Journal of Consulting and Clinical Psychology*. 1988; 56:885–892. doi:10.1037/0022-006X.56.6.885. [PubMed: 3204198]
- Howard KI, Kopta MS, Krause MS, Orlinsky DE. The dose-effect relationship in psychotherapy. *American Psychologist*. 1986; 41:159–164. doi:10.1037/0003-066X.41.2.159. [PubMed: 3516036]
- Howard KI, Lueger RJ, Maling MS, Martinovich Z. A phase model of psychotherapy outcome: Causal mediation of change. *Journal of Consulting and Clinical Psychology*. 1993; 61:678–685. doi: 10.1037/0022-006X.61.4.678. [PubMed: 8370864]
- Ilardi SS, Craighead W. The role of nonspecific factors in cognitive-behavior therapy for depression. *Clinical Psychology: Science and Practice*. 1994; 1:138–156.
- Jarrett RB, Kraft D, Doyle J, Foster BM, Eaves G, Silver PC. Preventing recurrent depression using cognitive therapy with and without a continuation phase. *Archives of General Psychiatry*. 2001; 58:381–388. doi:10.1001/archpsyc.58.4.381. [PubMed: 11296099]
- Jarrett, RB.; Minhajuddin, A.; Sanders, JM.; Friedman, E.; Thase, ME. Quantifying and qualifying the preventive effects of acute phase cognitive therapy: Pathways to personalizing care. (in preparation)
- Jarrett RB, Thase ME. Comparative efficacy and durability of continuation phase cognitive therapy for preventing recurrent depression: Design of a double-blinded fluoxetine- and pill placebo-controlled randomized trial with 2-year follow-up. *Contemporary Clinical Trials*. 2010; 31:355–377. [PubMed: 20451668]
- Jones BL, Nagin DS. Advances in group-based trajectory modeling and an SAS procedure for estimating them. *Sociological Methods & Research*. 2007; 35:542–571.
- Jung T, Wickrama KS. An introduction to latent class growth analysis and growth mixture modeling. *Social and Personality Psychology Compass*. 2008; 2:302–317. doi:10.1111/j.1751-9004.2007.00054.x.
- Kazdin AE. Mediators and mechanisms of change in psychotherapy research. *Annual Review of Clinical Psychology*. 2007; 3:1–27. doi:10.1146/annurev.clinpsy.3.022806.091432.
- Lutz W, Martinovich Z, Howard KI, Leon SC. Outcomes management, expected treatment response, and severity-adjusted provider profiling in outpatient psychotherapy. *Journal of Clinical Psychology*. 2002; 58:1291–1304. doi:10.1002/jclp.10070. [PubMed: 12357444]
- Otto MW, Teachman BA, Cohen LS, Soares CN, Vitonis AF, Harlow BL. Dysfunctional attitudes and episodes of major depression: Predictive validity and temporal stability in never-depressed, depressed, and recovered women. *Journal of Abnormal Psychology*. 2007; 116:475–483. doi: 10.1037/0021-843X.116.3.475. [PubMed: 17696703]

- Percevic R, Lambert MJ, Kordy H. What is the predictive value of responses to psychotherapy for its future course? Empirical explorations and consequences for outcome monitoring. *Psychotherapy Research*. 2006; 16:364–373. doi:10.1080/10503300500485524.
- Rush AJ, Gullion CM, Basco MR, Jarrett RB, Trivedi MH. The Inventory of Depressive Symptomatology (IDS): Psychometric Properties. *Psychological Medicine*. 1996; 26:477–486. doi:10.1017/S0033291700035558. [PubMed: 8733206]
- Schafer, JL. Analysis of incomplete multivariate data. Chapman & Hall; New York, NY: 1997.
- Tang TZ, DeRubeis RJ. Sudden gains and critical sessions in cognitive-behavioral therapy for depression. *Journal of Consulting and Clinical Psychology*. 1999; 67:894–904. doi: 10.1037/0022-006X.67.6.894. [PubMed: 10596511]
- Thase, ME.; Friedman, ES.; Minhajuddin, A.; Berman, S.; Houck, P.; Jarrett, RB. When is enough enough? Determining the minimum adequate trial of cognitive therapy for major depressive disorder. (in preparation)
- Uher RR, Muthén BB, Souery DD, Mors OO, Jaracz JJ, Placentino AA, McGuffin PP. Trajectories of change in depression severity during treatment with antidepressants. *Psychological Medicine*. 2010; 40:1367–1377. doi:10.1017/S0033291709991528. [PubMed: 19863842]
- Vittengl JR, Clark LA, Jarrett RB. Validity of sudden gains in acute phase treatment of depression. *Journal of Consulting and Clinical Psychology*. 2005; 73:173–182. doi:10.1037/0022-006X.73.1.173. [PubMed: 15709845]
- Vittengl JR, Clark LA, Kraft D, Jarrett RB. Multiple measures, methods, and moments: A factor-analytic investigation of change in depressive symptoms during acute-phase cognitive therapy. *Psychological Medicine*. 2005; 35:693–704. doi:10.1017/S0033291704004143. [PubMed: 15918346]
- Weissman MM, Bothwell S. Assessment of social adjustment by patient self-report. *Archives of General Psychiatry*. 1976; 33:1111–1115. [PubMed: 962494]
- Weissman MM. The Dysfunctional Attitudes Scale: A validation study. *Dissertation Abstracts International*. 1979; 40:1389B–1390B.
- Weissman MM, Olfson M, Gameroff MJ, Feder A, Fuentes M. A comparison of three scales for assessing social functioning in primary care. *American Journal of Psychiatry*. 2001; 158:460–466. doi:10.1176/appi.ajp.158.3.460. [PubMed: 11229989]
- Young, J.; Beck, AT. *Cognitive Therapy Scale: Rating Manual*. Center for Cognitive Therapy; Philadelphia, PA: 1980.

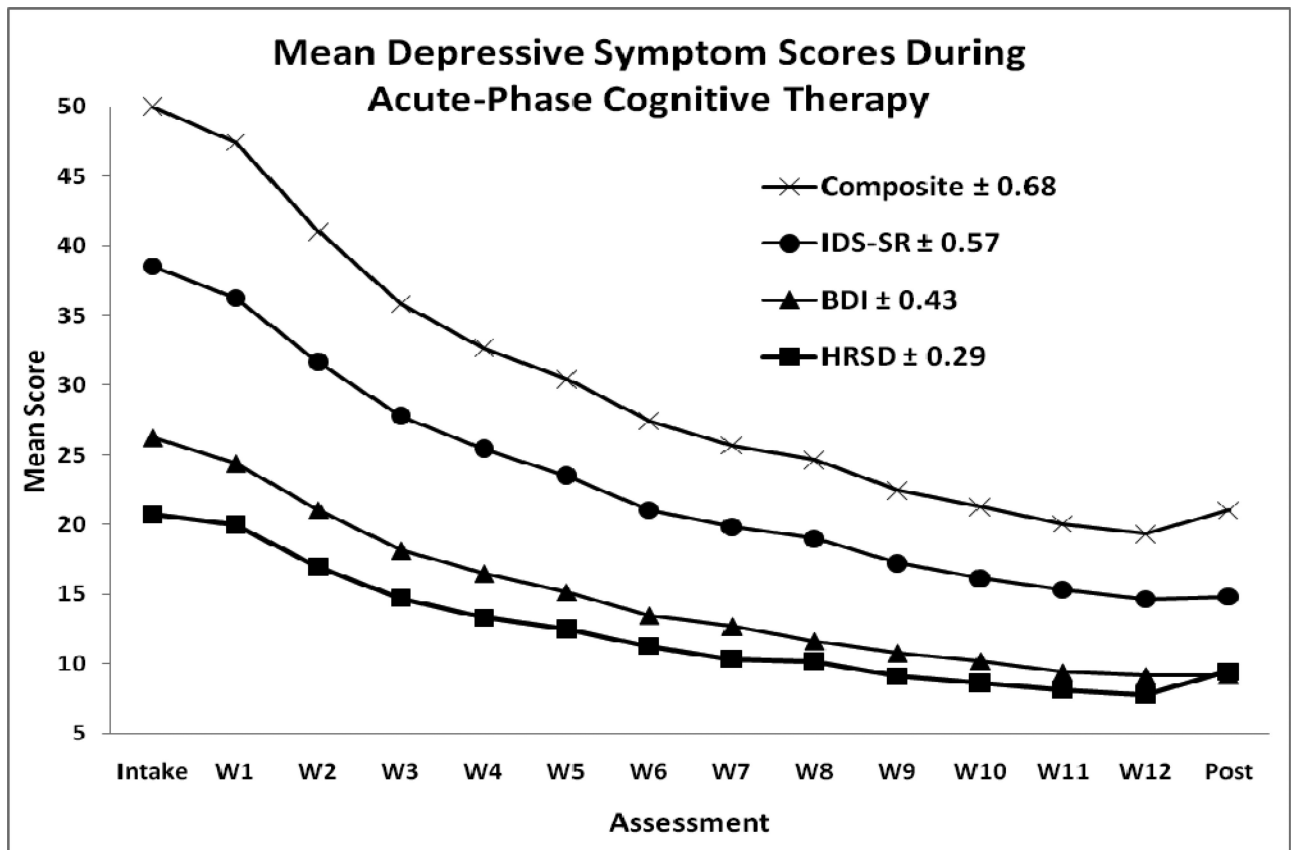


Figure 1.

Mean scores on three depressive symptom measures, and their standardized composite, by assessment. BDI = 21-item Beck Depression Inventory (21 items). HRSD = 17-item Hamilton Rating Scale for Depression. IDS-SR = 30-item Inventory for Depressive Symptomatology--Self-Report. *Intake* = first pre-treatment assessment. *W1-12* = assessment at cognitive therapy weeks 1-12. *Post* = post-treatment assessment.

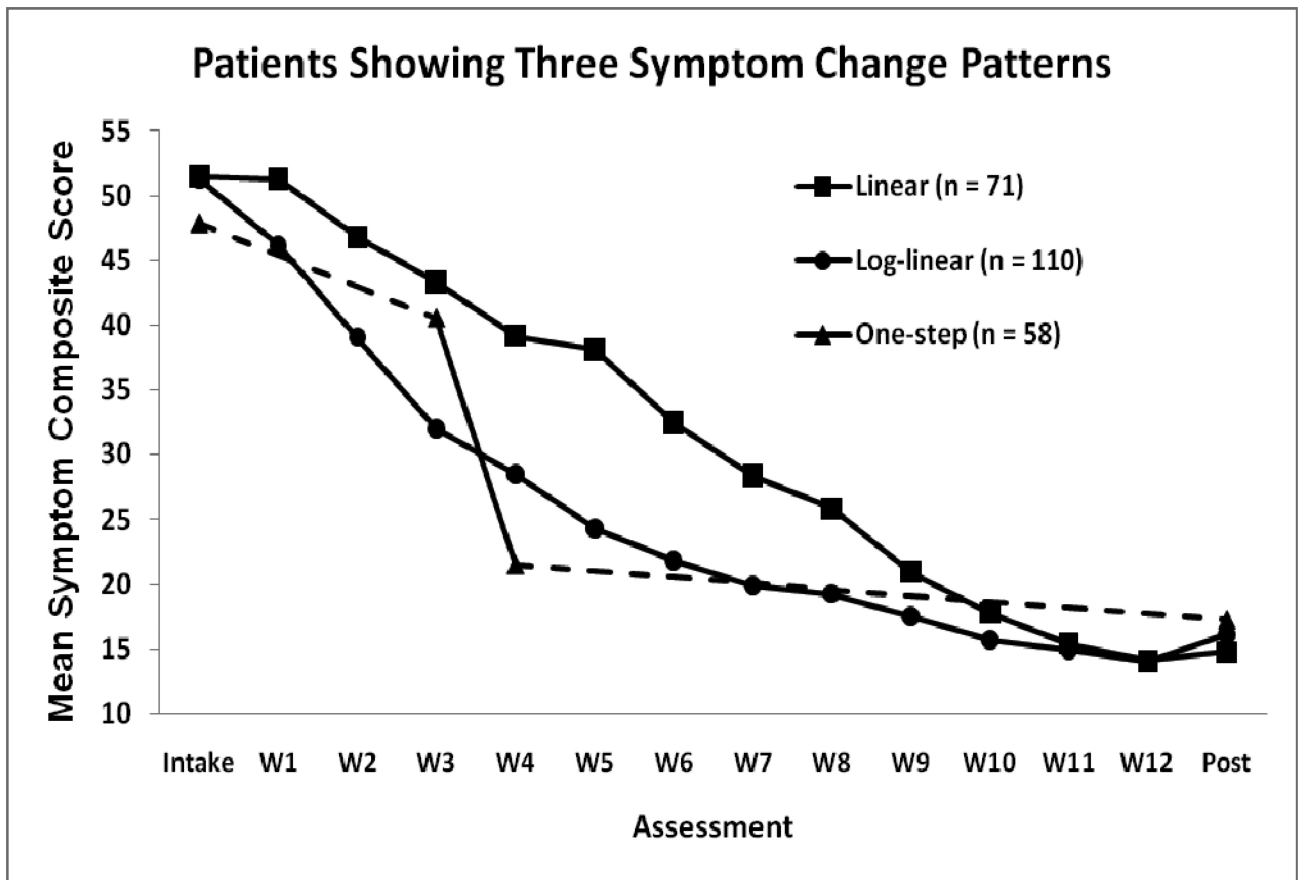


Figure 2. Mean symptom scores for groups of patients showing three a priori patterns of depressive symptom change. *Intake* = first pre-treatment assessment. *W1-12* = assessment at cognitive therapy weeks 1-12. *Post* = post-treatment assessment. The one-step line shows average *Intake* and *Post* scores, as well as average scores immediately before and after the median step between *W3-W4*; dotted lines indicate interpolation of remaining scores because the pre-/post-step durations vary among patients.

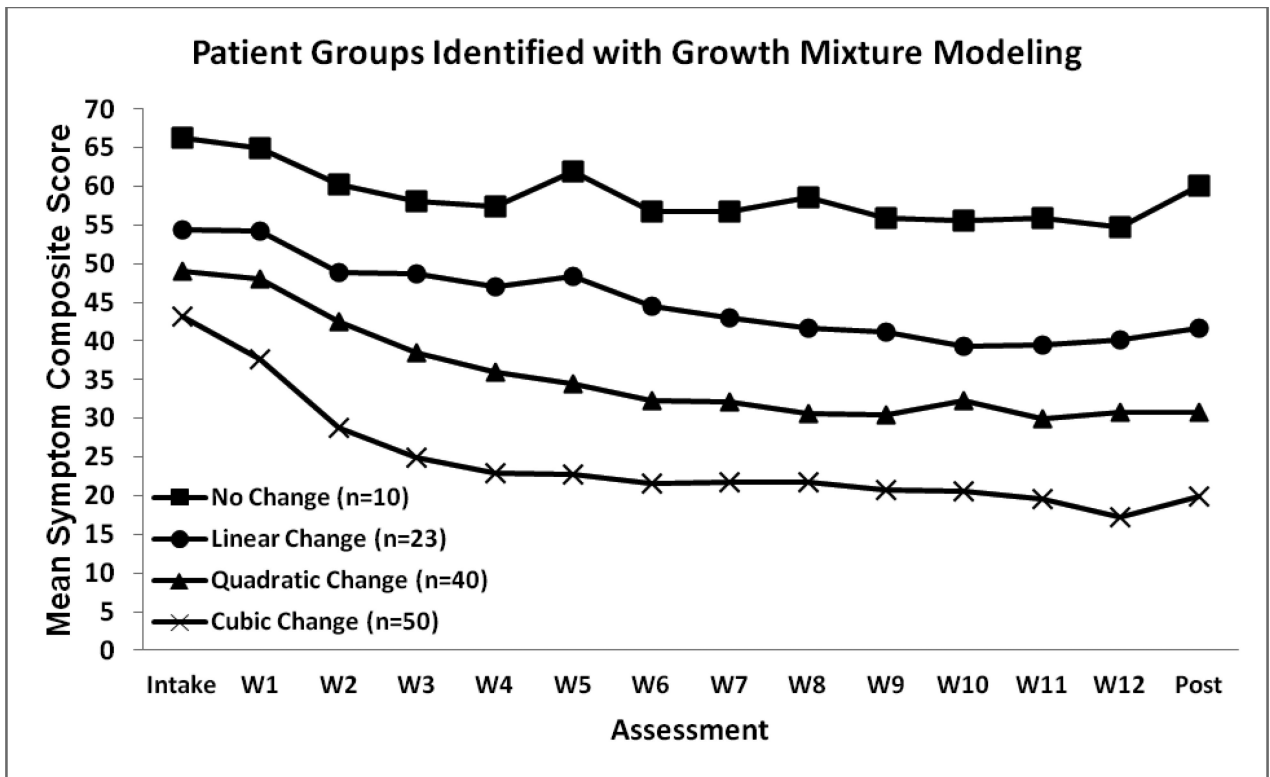


Figure 3.

Mean symptom scores for group of patients identified by growth mixture modeling. *Intake* = first pre-treatment assessment. *W1-12* = assessment at cognitive therapy weeks 1-12. *Post* = post-treatment assessment. All model coefficients significant at $p < .01$ in predicting symptoms y at assessment t numbered 1-14. No change group: $y_t = 58.83$. Linear change group: $y_t = 53.81 - 1.16t$. Quadratic change group: $y_t = 53.42 - 4.35t + 0.20t^2$. Cubic change group: $y_t = 52.50 - 10.31t + 1.10t^2 - 0.04t^3$.

Table 1

Varimax Rotated Loadings from Cross-Time Factor Analysis of Depressive Symptom Measures

Assessment	Factor 1 (zero-order)			Factor 1 (partial)			Factor 2 (zero-order)			Factor 2 (partial)		
	HRSD	BDI	IDSR	HRSD	BDI	IDSR	HRSD	BDI	IDSR	HRSD	BDI	IDSR
Intake	.13	.06	.11				.47	.69	.73			
Week 1	.20	.11	.13	.10	.02	.01	.67	.79	.80	.46	.51	.53
Week 2	.19	.15	.17	.06	.04	.03	.69	.83	.84	.62	.69	.74
Week 3	.37	.30	.33	.26	.23	.24	.67	.76	.79	.72	.70	.76
Week 4	.47	.45	.45	.38	.39	.38	.62	.68	.72	.66	.70	.75
Week 5	.53	.55	.58	.45	.50	.53	.56	.63	.64	.62	.63	.65
Week 6	.63	.62	.65	.56	.60	.62	.48	.57	.59	.57	.56	.59
Week 7	.70	.70	.71	.67	.70	.70	.41	.47	.47	.42	.42	.44
Week 8	.74	.79	.80	.70	.77	.77	.32	.39	.37	.35	.38	.39
Week 9	.81	.84	.86	.78	.84	.85	.18	.27	.29	.23	.22	.27
Week 10	.80	.86	.86	.79	.87	.87	.15	.24	.27	.15	.17	.20
Week 11	.85	.90	.90	.84	.92	.90	.16	.18	.21	.15	.12	.18
Week 12	.82	.86	.89	.81	.88	.89	.15	.19	.21	.15	.13	.17
Post	.80	.82	.86	.79	.83	.86	.20	.22	.23	.17	.13	.17

Note. N = 362. Loadings .50 in bold type. BDI = 21-item Beck Depression Inventory. HRSD = 17-item Hamilton Rating Scale for Depression. IDS-SR = 30-item Inventory for Depressive Symptomatology--Self-Report. Intake = first pre-treatment assessments. Week 1-12 = assessment at cognitive therapy weeks 1-12. Post = post-treatment assessment. Factor analyses of zero-order and partial correlations (controlling Intake) test interrelations of symptom severity and change, respectively

Table 2

Formulas for Pairwise Conversions among Three Depressive Symptom Measures Based on Their Common Factor

Measure Estimated ($\pm SE$)	Measure Known		
	BDI	HRSD	IDS-SR
BDI (± 2.65) =	---	$-3.86 + 1.51(\text{HRSD})$	$-1.75 + 0.72(\text{IDS-SR})$
HRSD (± 2.92) =	$2.55 + 0.66(\text{BDI})$	---	$1.40 + 0.48(\text{IDS-SR})$
IDS-SR (± 2.43) =	$2.42 + 1.38(\text{BDI})$	$-2.92 + 2.09(\text{HRSD})$	---

Note. BDI = 21-item Beck Depression Inventory. HRSD = 17-item Hamilton Rating Scale for Depression. IDS-SR = 30-item Inventory for Depressive Symptomatology--Self-Report. Observed ranges: BDI 0-59, HRSD 0-37, IDS-SR 0-74. To convert an HRSD score of 20 to the IDS-SR, for example, use the formula $\text{IDS-SR} = -2.92 + 2.09(\text{HRSD})$; substitute the known HRSD score in the equation, $-2.92 + 2.09(20)$; and solve to yield $\text{IDS-SR} \approx 39$.

Table 3

Multilevel Models Predicting Depressive Symptom Composite Scores across 14 Assessments

<u>Fixed Effect B(SE)</u>	<u>Linear Model</u>	<u>Log-linear Model</u>
Intercept	47.21 (0.58) *	53.33 (0.59) *
Assessment slope	-2.30 (0.058) *	-29.93 (0.68) *
<u>Random Effects Variance Estimate (SE)</u>		
Intercept	104.95 (8.97) *	104.33 (9.34) *
Assessment slope	0.99 (0.090) *	134.59 (12.24) *
<u>Model Fit</u>		
Bayes Information Criterion	36495	35851

Note. $N = 362$. Composite scores are the standardized average of the 21-item Beck Depression Inventory, 17-item Hamilton Rating Scale for Depression, and 30-item Inventory for Depressive Symptomatology--Self-Report. Assessment slopes computed with assessment coded 1-14 in linear models and log(assessment) in the log-linear models. Smaller Bayes Information Criterion values indicate better model fit.

* $p < .01$.

Table 4
Fit of Individual Regression Models to Patients in Three Symptom Change Pattern Groups

Change Pattern	n (%)	Model Fit R^2			Intercept			Slope Coefficient		
		M	SD	Min. Max.	M	SD	Min. Max.	M	SD	Min. Max.
Linear	71 (20%)	.87	.05	.76 .97	55.79	10.50	36.15 81.71	-3.25	0.99	-5.44 -1.05
Log-linear	110 (30%)	.86	.06	.75 .95	53.56	10.99	32.76 77.04	-35.58	10.54	-62.33 -15.95
One-step	58 (16%)	.88	.04	.81 .96	44.99	8.92	25.98 68.07	-27.84	8.61	-43.74 -14.48
Unclassified	123 (34%)	---	---	---	---	---	---	---	---	---

Note. Regressions predicted a series of 14 depression symptom scores with linear, log-linear, and one-step functions. Patients were assigned to the pattern showing the best fit compared to cases in a simulation study with no average symptom change. Minimum R^2 values were at the 95th percentile of simulated cases. Unclassified patients had fits below the 95th percentile of simulated cases.

Table 5

Differentiation of Patients Showing Four Patterns of Change in Depressive Symptom Scores

Test Variable	Linear		Log-linear		One-step		Unclassified	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Demographics at Intake								
Age	41.79	12.82	42.29	12.83	45.34	10.99	44.42	11.68
Women	69%		75%		57%		69%	
Ethnicity (exact $p = .15$ over 6 ethnic groups)								
Illness Characteristics at Intake								
Age of MDD Onset	21.68	11.52	20.10	9.78	20.59	11.69	22.15	10.17
Years of MDD	19.58	11.15	21.63	12.40	24.33	12.75	21.84	11.39
MDE Months	20.85	31.08	25.11	52.66	20.64	24.65	34.46	57.65
Depressive Symptom Severity								
Composite at Intake	51.48	9.58	51.25	9.29	47.84	9.26	49.05	10.97
Composite at Exit *	14.76 ^a	9.06	16.11 ^a	9.74	17.24 ^a	11.97	30.81 ^b	15.54
Treatment Response *	93% ^a		87% ^a		88% ^a		45% ^b	
Stable Remission *	8% ^a		21% ^b		26% ^b		1% ^c	
Social Functioning								
SAS-SR at Intake	2.64	0.40	2.55	0.44	2.50	0.39	2.56	0.46
SAS-SR at Exit *	1.85 ^a	0.36	1.89 ^a	0.40	1.88 ^a	0.41	2.18 ^b	0.44
IIP at Intake	1.77	0.45	1.70	0.47	1.66	0.64	1.60	0.54
IIP at Exit *	1.06 ^a	0.43	1.04 ^a	0.56	1.10 ^a	0.60	1.33 ^b	0.59
Cognitive Content								
DAS at Intake	151.52	33.62	156.21	34.79	144.44	33.56	148.79	34.85
DAS at Exit *	107.12 ^a	27.24	114.42 ^a	36.61	106.68 ^a	33.33	126.38 ^b	29.69
BHS at Intake	12.24	5.50	12.28	5.08	12.05	4.87	11.77	4.88
BHS at Exit *	3.67 ^a	3.59	5.23 ^b	4.74	4.98 ^{ab}	4.39	8.51 ^c	5.65

* $p < .05$ differences among groups; within a row, means with different superscripts vary. MDD = major depressive disorder. MDE months = duration of current major depressive episode. Symptom severity is a standardized composite of the Hamilton Rating Scale for Depression, Inventory for Depressive Symptomatology--Self-Report, and Beck Depression Inventory. SAS-SR = Social Adjustment Scale--Self-Report. IIP = Inventory of Interpersonal Problems. DAD = Dysfunctional Attitudes Schedule. BHS = Beck Hopelessness Scale. Response = no MDE and 17-item Hamilton Rating Scale for Depression ≥ 12 at exit. Stable remission = no MDE at exit and the last seven scores on the 17-item Hamilton Rating Scale for Depression ≤ 6 .