Randomized Trial of Behavioral Activation, Cognitive Therapy, and Antidepressant Medication in the Prevention of Relapse and Recurrence in Major Depression

Keith S. Dobson
University of Calgary

Steven D. Hollon
Vanderbilt University

Sona Dimidjian
University of Colorado

Karen B. Schmaling
University of North Carolina at Charlotte

Robert J. Kohlenberg
University of Washington

Robert J. Gallop
West Chester University

Shireen L. Rizvi
New School for Social Research

Jackie K. Gollan
University of Chicago

David L. Dunner and Neil S. Jacobson
University of Washington

This study followed treatment responders from a randomized controlled trial of adults with major depression. Patients treated with medication but withdrawn onto pill-placebo had more relapse through 1 year of follow-up compared to patients who received prior behavioral activation, prior cognitive therapy, or continued medication. Prior psychotherapy was also superior to medication withdrawal in the prevention of recurrence across the 2nd year of follow-up. Specific comparisons indicated that patients previously exposed to cognitive therapy were significantly less likely to relapse following treatment termination than patients withdrawn from medication, and patients previously exposed to behavioral activation did almost as well relative to patients withdrawn from medication, although the difference was not significantly different. Differences between behavioral activation and cognitive therapy were small in magnitude and not significantly different across the full 2-year follow-up, and each therapy was at least as efficacious as the continuation of medication. These findings suggest that behavioral activation may be nearly as enduring as cognitive therapy and that both psychotherapies are less expensive and longer lasting alternatives to medication in the treatment of depression.

Keywords: behavioral activation, cognitive therapy, antidepressant medication, major depression, relapse

Antidepressant medication (ADM) has been shown to prevent the return of symptoms associated with major depression for as long as it is continued or maintained (American Psychiatric Association, 2000). However, there is little evidence that having taken medication does anything to alter the risk factors that lead to subsequent relapse and recurrence, and most patients with chronic or recurrent depression are encouraged to stay on medication indefinitely (Hollon, Thase, & Markowitz, 2002). Further, there is a recognized need for alternatives to medications, given their potential for side effects and some patients' preferences for non-pharmacological treatments for depression (Hollon & Shelton, 2001; van Schaik et al., 2004).

In contrast to ADM, cognitive therapy (CT) has been demonstrated to have an enduring effect that protects against relapse and possibly recurrence (Hollon, Stewart, & Strunk, 2006; Kovacs, Rush, Beck, & Hollon, 1981; Simons, Murphy, Levine, & Wetzel, 1986). For example, a meta-analysis of the effects of CT concluded that the 1-year relapse rate for patients previously treated with CT was approximately 25%, compared to 50% for patients previously treated with ADM (Gloaguen, Cottraux, Cucherat, & Blackburn, 1998). In fact, two trials have now demonstrated that the preventative effect of prior CT is comparable in magnitude to that produced by keeping patients on ADM (Evans et al., 1992; Hollon et al., 2005). It is important that the cost estimates in the Hollon et al. (2005) study favored prior CT by the time of 1 year of follow-up, suggesting that CT is not only as clinically efficacious as ADM but also more cost effective in the longer term (see also Antonuccio, Thomas, & Danton, 1997).

Although there is a considerable literature regarding the enduring effects of prior CT, less is known about the relative long-term efficacy of other psychosocial treatments. In this regard, it is important to evaluate the potential enduring effects of behavioral activation (BA; Jacobson, Martell, & Dimidjian, 2001; Martell, Addis, & Jacobson, 2001). A component analysis of CT suggested that an earlier version of BA was as efficacious as CT during acute treatment (Jacobson et al., 1996) and no less enduring following treatment termination (Gortner, Gollan, Dobson, & Jacobson, 1998). In a subsequent trial, an expanded BA model was found to be comparable to ADM with respect to the reduction of acute distress regardless of level of initial severity and superior to CT among more severely depressed patients (Dimidjian et al., 2006).

The primary aim of this study was to determine the enduring effects of prior exposure to BA, prior exposure to CT, and continued treatment with ADM in the context of a placebo-controlled trial. We examined enduring effects relative to both relapse (the return of the treated episode of depression) and recurrence (the onset of a new episode of depression) using the consensus definitions established by the MacArthur Group (Frank et al., 1991). We anticipated that both prior BA and prior CT would have enduring effects in the prevention of relapse that would be comparable to continued medication and superior to medication withdrawal during a 1-year follow-up phase following acute treatment. Additionally, we investigated whether the effects of prior psychosocial treatments extended beyond the prevention of relapse to the prevention of recurrence during a 2nd-year follow-up after acute treatment. During this 2nd year, patients previously continued on ADM were withdrawn from medication. The primary predictions during the 2nd-year follow-up phase were that both prior BA and prior CT would result in reduced rates of recurrence relative to

ADM withdrawal. Finally, the direct treatment costs for psychotherapy and medication were compared.

Method

Participants

The participants for this study consisted of adult outpatients who responded to acute phase treatment for depression from the Dimidjian et al. (2006) study, based in Seattle, WA. Participants in that study met criteria for *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.; American Psychiatric Association, 1994) major depressive disorder on the basis of diagnostic interviews, had scores of 20 or above on the Beck Depression Inventory II (BDI-II; Beck, Steer, & Brown, 1996) and scores of 14 or above on the 17-item version of the Hamilton Rating Scale for Depression (HRSD; Hamilton, 1960), and did not meet criteria for a number of common exclusion diagnoses in depression psychotherapy studies. The University of Washington institutional review board approved the protocol.[1] All participants provided written informed consent prior to enrollment in the study.

We randomly assigned eligible participants to an active acute phase treatment condition using a computer-generated list (a subset of patients also were assigned to pill-placebo during the acute phase but are not considered in the current report). These treatments were 16 weeks of ADM (paroxetine; $n = 100$), CT ($n = 45$), or BA ($n = 43$). Approximately twice as many patients were randomized to ADM compared to each of the psychotherapies, which allowed half to be continued on medication and half to be withdrawn onto pill-placebo at the beginning of the 1st year of follow-up. Therapists were trained and supervised on an ongoing basis to ensure that the treatment protocols were enacted as planned. These treatment protocols included the standard manual for cognitive therapy (Beck, Rush, Shaw, & Emery, 1979), as well as the published manual for behavioral activation (Martell et al., 2001). Adherence ratings indicated that BA consisted predominantly of behavioral strategies and CT of a mixture of behavioral and cognitive strategies. Competence ratings in the CT condition were in the acceptable range but were not collected in BA due to the lack of a competence measure (Dimidjian et al., 2006). ADM was conducted according to guidelines commonly used in such trials (Fawcett, Epstein, Fiester, Elkin, & Autry, 1987). The maximum dosage of paroxetine was 50 mg/day, with dosages adjusted according to a predetermined regimen up to a maximally tolerated dosage. Of the patients originally assigned to an active treatment, 106 patients no longer met diagnostic criteria for major depressive disorder at the end of the acute phase of treatment and formed the sample for the current study.

Measures

Measures of depressive severity and relapse/recurrence. The BDI-II (Beck et al., 1996) is the most commonly used self-report measure of clinical depression severity. It consists of 21 items,

---

which yield a range of scores from 0–63. The psychometric criteria for the BDI-II are generally held to be excellent when used with outpatients (Beck et al., 1996; Nezu, Ronan, Meadows, & McClure, 2000). The 17-item version of the HRSD (Hamilton, 1960) is the most commonly used interview measure of depression and is widely employed to assess depression severity in psychotherapy and antidepressant medication research (Nezu et al., 2000). The Longitudinal Interview for Follow-up Evaluation (LIFE; Keller et al., 1987) is a semistructured interview commonly employed to retrospectively assess the longitudinal course of various psychiatric disorders, including major depression. Psychiatric status ratings (PSRs) were obtained retrospectively for each study week and ranged from 1 (*absence of symptomatology*) to 6 (*definite and severe presence of symptomatology*) for major depressive disorder. In addition, the LIFE was used to obtain information about the initiation of other treatments, with special attention to those for depression, which could have a significant effect on the estimates of the long-term benefits of treatment.

Definitions of relapse and recurrence. Rates of relapse and recurrence served as the primary outcome measures in this study. *Relapse* was defined as either HRSD scores of 14 or greater or PSRs of 5 or greater for 2 successive weeks during the 1st year of follow-up, whereas *recurrence* was defined with the same criteria during the 2nd year of follow-up. These definitions are consistent with the MacArthur consensus criteria (Frank et al., 1991) and reflect the notion that recently remitted patients are at elevated risk for a relapse into the symptoms associated with the treated episode, whereas patients who have gone for a longer period of time have recovered from the index episode and any subsequent symptom return represents the onset of a wholly new episode, or recurrence (Hollon et al., 2002).

Reliability of evaluators. Independent evaluators were trained, monitored, and certified to conduct study evaluations. Raters were blind to treatment condition and supervised weekly to prevent rater drift. A randomly selected subset of taped follow-up interviews was rated by one of two criterion judges. Intraclass correlation coefficients were computed between the original evaluator's rating and the criterion judge's (Shrout & Fleiss, 1979). Results indicated reliability of .98 for the HRSD ($n = 27$) and .79 for the weekly PSR depression ratings ($n = 25$).

Procedure

Participants were followed to the point of relapse or recurrence for up to 2 years following response to acute treatment. The 1st year of follow-up compared four groups of acute phase treatment responders: those who received prior CT ($n = 30$), those who received prior BA ($n = 27$), and those who received prior ADM ($n = 49$), with the latter divided into two groups as specified below. Patients who had received prior BA or prior CT received no further treatment during any follow-up phase of the study. Treatment responders who had received ADM were randomized by previous assignment, managed by the participant coordinator, and were continued on active medication (cADM; $n = 28$) or withdrawn onto pill-placebo at the beginning of the 1st year follow-up phase according to a predetermined 2-week taper schedule (cPLA; $n = 21$). Patients in cADM and cPLA continued to see their pharmacotherapists biweekly for the first 2 months and monthly thereafter during the rest of the 1st year of follow-up, and clini-

cians treated them using standard clinical management strategies. Pharmacotherapists, evaluators, and patients were blind to medication status, meaning that the pharmacological portion of the trial was conducted triple blind. cADM participants were continued on the same dosage prescribed during acute treatment, with adjustments allowed as clinically indicated. At the end of the 1st year of follow-up, pharmacotherapists discontinued cADM patients' medication using the same taper schedule as was used for the cPLA participants. Patients in both cADM and cPLA were seen biweekly during the taper period and were then assessed during the 2nd follow-up year. Participants who relapsed following withdrawal onto pill-placebo were provided with free treatment by study therapists; patients who relapsed in the other conditions were provided with referrals to appropriate community agencies.

Participants completed the outcome assessment instruments biweekly for the first 2 months of the 1st-year follow-up phase, at Month 3, and again at Months 6 and 12. Assessments also were conducted at Months 13, 14, 18, and 24 during the 2nd follow-up year. Additional ad hoc assessments were conducted whenever a new episode of depression was suspected, on the basis of elevated HRSD scores or reports of the participant or treating pharmacotherapist.

Data Analysis

Analysis of variance (ANOVA) was used to test for differences among groups on continuous variables. Chi-square tests of independence were used for categorical variables, although the chi-square test was replaced by Fisher's exact test for small or empty cells. We first compared patients who entered this study to those from the original sample who did not complete or failed to respond to acute treatment in order to assess the generalizability or external validity of the findings (see Table 1 for a comparison of those participants who entered the current follow-up study to the total study sample). We next used the same sets of analyses to compare the treatment groups at the beginning of this follow-up study to ensure the differential retention did not bias the internal validity of the study or the causal inferences that could be drawn. Differential retention can systematically undermine the comparability of groups achieved by randomization at the start of the trial; if patients at differing levels of risk for relapse or recurrence are more or less likely to complete or respond to the respective conditions, then acute treatment can act as a "differential sieve" that systematically undermines the equivalence produced by initial randomization (Klein, 1996). Given their cautionary role, p values were set at $p = .10$ for these analyses.

We estimated survival curves and rates of relapse and recurrence across the follow-up using the Cox proportional hazards model (Collett, 1994; Cox & Oakes, 1984). We used the log-rank test to compare survival rates across treatment conditions and conducted specific contrasts using the Cox regression for the survival curves. Statistical significance was set at $p < .05$ (two-tailed). Specific comparisons were conducted between the active interventions (prior CT [pCT], prior BA [pBA], and cADM) versus medication withdrawal (cPLA) across the 1st follow-up year with respect to relapse. We applied similar analyses to the data collected during the 2nd follow-up year to examine rates of subsequent recurrence, with the sample restricted to all recovered patients who survived the 1st follow-up year without relapse.

Table 1
Demographic and Clinical Characteristics of the Sample

Variable	Prior BA ($n = 27$)	Prior CT ($n = 30$)	ADM ($n = 49$)	Follow-up sample ($n = 106$)	Total sample ($N = 241$)
Demographic characteristics					
Mean age (*SD*)	41.54 (11.93)	39.01 (11.06)	38.93 (10.04)	40.04 (10.85)	39.90 (10.97)
Gender					
Female	46.4	73.3	78.2	68.9	66.0
Male	53.6	26.7	21.8	31.1	34.0
Race/ethnicity					
Caucasian	85.7	86.7	80.0	82.1	81.7
Minority	14.3	13.3	20.0	17.9	18.3
Married or live with a partner	46.4	41.0	36.3	40.6	39.0
Have 1 or more children	46.4	53.3	43.6	48.1	45.2
College graduate	45.5	66.4	63.8	59.5	58.7
Characteristics of depression subtype and history					
Mean age at first onset (*SD*)	29.50 (15.58)	26.00 (14.26)	27.02 (11.94)	27.59 (13.59)	27.65 (13.27)
Mean no. of prior episodes (*SD*)	1.00 (1.25)	1.30 (1.39)	1.12 (1.30)	1.11 (1.30)	1.00 (1.44)
Chronicity, episode > 2 years duration	26.3	27.6	46.1	32.1	34.4
Recurrent subtype	53.6	66.7	56.4	58.5	57.7
Melancholic subtype	25.0	26.7	20.0	21.7	30.3
Atypical subtype	10.7	13.3	14.5	13.2	17.4

Note. Data are percentages unless otherwise noted. BA = behavioral activation; CT = cognitive therapy; ADM = antidepressant medication.

These analyses compared the two psychosocial interventions (pCT and pBA) with medication withdrawal at the start of that period (cADM). Finally, the two psychosocial conditions (pCT and pBA) were directly compared across the full 2-year follow-up, because any preventive effects associated with either should be robust with respect to both relapse and recurrence.

Results

Patient Flow and Attrition

We followed up with all patients for 24 months, unless the original episode of depression returned (relapse), a new episode started (recurrence), or a participant wished to discontinue study participation, moved away, or was unavailable for follow-up evaluations (see Figure 1). Patients who were lost to follow-up or who returned to treatment for depression in the absence of a documented relapse were credited with the amount of time that they survived without relapse or recurrence, up until the point that their data were censored due to attrition (Kaplan & Meier, 1958). Data were available to estimate risk to the point of relapse or recurrence for 92 of the 106 patients (86.8%) who entered the follow-up period. Most attrition occurred early in the 1st year of follow-up, consisting of 3 participants from prior CT, 5 from prior BA, 2 from cADM, and 4 from cPLA. Further, data from 1 participant each from prior BA and CT were censored due to receiving antidepressant medications outside of the study in the 1st year of follow-up.

Sample Characteristics

Treatment responders who entered the follow-up did not differ significantly from those patients who failed to complete or respond to acute treatment on any of the characteristics listed in Table 1. The average age of participants was approximately 40 years in a sample that was predominantly female and Caucasian. Slightly less than half of the sample was married or living with a partner, and approximately half had children and were college graduates. Slightly more than half the sample met criteria for recurrent depression, and participants had, on average, 1.11 prior episodes of depression. About one third of the sample met the criteria for chronic depression (more than 2 years in duration) for the treated episode, with melancholic and atypical subtypes being somewhat less common (21.7% and 14.2%, respectively).

Few of the indexes explored in Table 1 met criteria for possible confounds that could have biased the subsequent survival analyses. Gender, which had differentiated treatment groups at the outset of the acute phase (Dimidjian et al., 2006), also differentiated among the treatment groups at the end of the acute phase of treatment, $\chi^2(2, N = 106) = 10.42$, $p < .01$. Residual depression, as measured by the BDI-II at posttest, also was significantly different between the prior CT ($M = 9.57$, $SD = 8.74$), prior BA ($M = 6.59$, $SD = 6.35$), and ADM ($M = 5.61$, $SD = 5.05$) conditions, $F(2, 103) = 3.40$, $p = .037$, and was a significant predictor in survival analyses, as higher residual levels of depression were associated with lower levels of survival.

We had some concerns about using posttreatment BDI-II scores as a covariate in the survival analyses because such an adjustment could create a bias in favor of the group with higher residual scores, which in this case was CT. At the same time, adjusting for a significant predictor of subsequent outcomes should reduce variance in outcomes, thereby enhancing design power. Analyses were conducted both ways, with no evidence of any bias in favor of CT, but with considerable gain in precision when posttreatment BDI-II scores were used as a covariate. Given these considerations and the patterns observed in the data, we conducted survival analyses using both gender and posttreatment BDI-II scores as covariates.
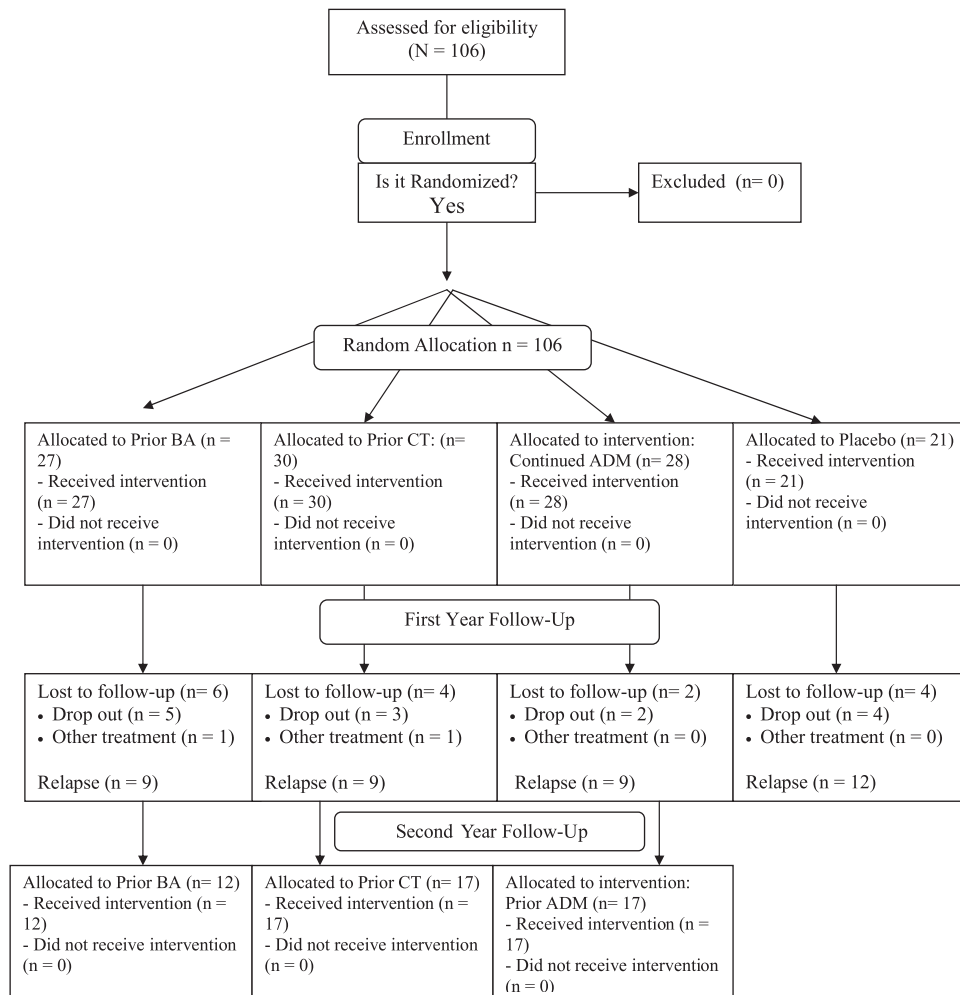
Figure 1. CONSORT flow chart. BA = behavioral activation; CT = cognitive therapy; ADM = antidepressant medication.

Prevention of Relapse

In order to give a picture of overall rates of relapse and recurrence, Figure 2 presents the survival curves across the entire 2-year follow-up phase of the study, but we focus first on rates of relapse across the 1st year of follow-up. As can be seen in Figure 2, relapse was especially likely to occur at the start of the 1st follow-up year, especially for medication responders withdrawn onto pill-placebo (cPLA). The rates of relapse during the 1st follow-up year, estimated from the Cox regression analysis, were 39% for prior CT, 50% for prior BA, 53% for cADM, and 59% for medication withdrawal (cPLA). When considered as a set, the active treatments (prior CT, prior BA, and cADM) were superior to withdrawal onto pill-placebo (cPLA), $\chi^2(1, N = 106) = 4.07$, $p = .04$. When considered separately, prior CT was significantly superior to cPLA, $\chi^2(1, N = 51) = 5.30$, $p = .02$, and prior BA demonstrated a nonsignificant trend, $\chi^2(1, N = 51) = 2.81$, $p = .09$, but cADM was not significantly different from cPLA, $\chi^2(1, N = 49) = 0.97$, $p = .33$.

As an index of the effect sizes for the treatment groups, hazard ratios were calculated between cPLA and each of the respective active treatments. Prior exposure to CT was associated with a hazard ratio of .36 relative to cPLA. The relative reduction of risk, computed by subtracting the hazard ratio from 1.0, indicated that prior exposure to CT reduced the risk for relapse by 64% relative to medication withdrawal. This is quite similar to the 70% reduction in risk shown in another recent trial (Hollon et al., 2005). cADM was associated with a hazard ratio of .67 relative to cPLA, or a reduction in risk of about 33%, which was somewhat lower than the 50% reduction observed in that same earlier trial by Hollon et al. (2005). Prior exposure to BA was associated with a hazard ratio of .49 and a reduction in risk of 51%. This is only slightly lower than the reduction of risk observed for prior CT in both the current study and the earlier study by Hollon et al. (2005) and is at least comparable to the magnitude of the effect typically observed for the continuation of medication (Hollon et al., 2002).

Prevention of Recurrence

We next considered whether the enduring effects of psychotherapy extended to the prevention of recurrence in recovered patients. These comparisons focus on data from the second half of Figure 2,
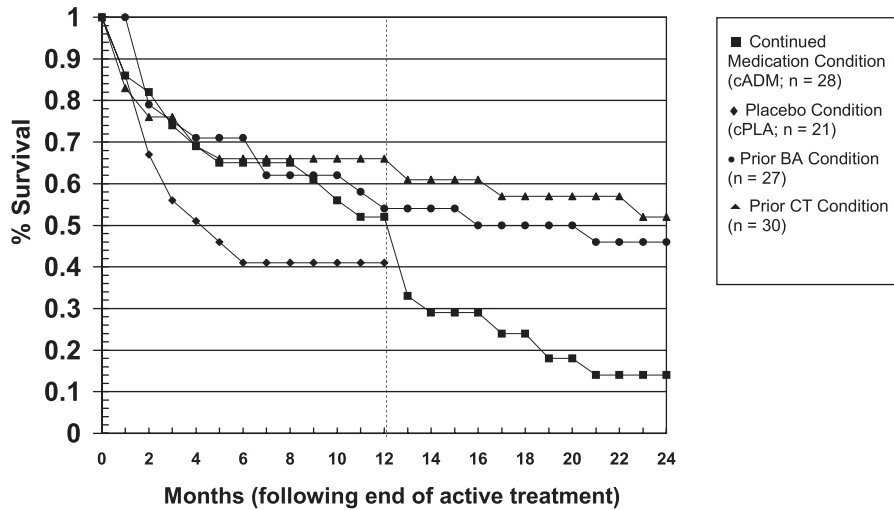
Figure 2. Cumulative proportion of treatment responders who survived without relapse over the 2 years of follow-up (the vertical line indicates the point at which patients on continuation medication were withdrawn from medication). BA = behavioral activation; CT = cognitive therapy.

but for the sample of 48 patients who survived the 1st follow-up year without relapse. Patients in cADM were withdrawn from medication at the beginning of that period. As illustrated in Figure 3, the enduring effect of prior psychotherapy extended to the prevention of recurrence. The rates of recurrence during the 2nd follow-up year, estimated from the Cox regression analysis, were 24% for prior CT, 26% for prior BA, and 52% for prior cADM. The effect of prior psychotherapy (prior CT and prior BA) evidenced a nonsignificant trend compared to the effect of prior cADM, $\chi^2(1, N = 46) = 3.58, p = .06$. The specific comparisons between prior CT and prior cADM, $\chi^2(1, N = 34) = 2.71, p = .10$, and between prior BA and prior cADM, $\chi^2(1, N = 29) = 1.40, p = .24$, were not significant.

As an index of the effect size of these differential recurrence rates, prior exposure to either CT or BA was associated with a hazard ratio of .37 relative to prior cADM, meaning that both CT and BA reduced the risk of recurrence by about 63% relative to

medication withdrawal. This result is somewhat lower than the 85% reduction in risk for recurrence shown by prior exposure to CT relative to medication withdrawal following in the study by Hollon et al. (2005).

The survival analyses conducted in the 2nd year of follow-up provided an opportunity to replicate the enduring effects of prior psychotherapy relative to medication withdrawal, but it was somewhat underpowered because it was restricted to only that subset of the sample that showed full recovery. We therefore repeated the survival analysis across the entire 2-year follow-up period (excluding the cPLA group). This comparison tested the long-term effects of prior BA and prior CT, relative to a group of patients who were continued on medication for the 1st year and then withdrawn from medication in the 2nd year (cADM). Although pharmacotherapists often keep their patients on medications indefinitely, many do not, and this specific comparison addresses the enduring effects of the two psychosocial interventions relative to patients who were suc-
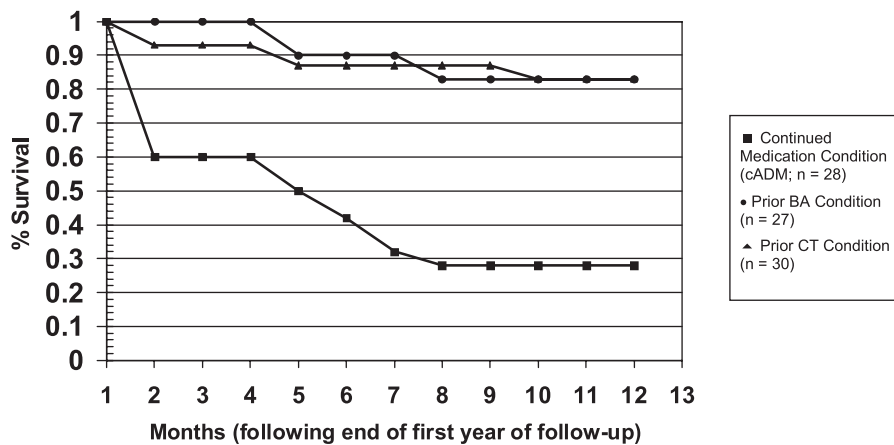

Figure 3. Cumulative proportion of recovered patients who survived without recurrence during the 2nd year of follow-up. BA = behavioral activation; CT = cognitive therapy.

cessfully treated with medication and then withdrawn from that medication. Further, this comparison provided a potential replication of the analyses conducted in the 1st year of follow-up, in which prior psychotherapy was compared to medication withdrawal.

The overall effect for treatment was significant, $\chi^2(2, N = 85) = 6.59$, $p = .04$, as the two psychosocial treatments (prior CT and prior BA) were superior to the continuation of medication followed by medication withdrawal (cADM). When considered separately, prior CT was significantly superior to cADM, $\chi^2(1, N = 58) = 5.97$, $p = .02$, whereas prior BA exhibited a nonsignificant trend in the same direction, $\chi^2(1, N = 55) = 3.09$, $p = .08$. When considered from the perspective of hazard ratios and relative reduction in risk of recurrence, prior exposure to BA was associated with a hazard ratio of .53 and reduction in risk of 47% relative to cADM. Exposure to prior CT was associated with a hazard ratio of .42 relative to cADM, or a reduced risk of recurrence of 58%. As can be seen in Figure 2, the bulk of the differences occurred when cADM patients were withdrawn from medication at the beginning of Year 2.

Because CT has demonstrated enduring treatment effects (Hollon et al., 2006), and as it typically provided a somewhat larger advantage relative to medication withdrawal compared to BA, the two psychotherapies were directly compared with the maximal power provided by the full 2-year comparison. CT and BA did not significantly differ, $\chi^2(1, N = 57) = 0.32$, $p = .57$. Exposure to prior CT was associated with a hazard ratio of .73 relative to prior BA, or a reduction in risk of 27%.

Sustained Response and Recovery

Although survival analyses often are used to examine rates of relapse and recurrence, they are susceptible to bias due to differential retention, because only those patients who both complete and respond to acute treatment are included in the former and only recovered patients who stay free from relapse are included in the latter. As a consequence, we also examined rates of *sustained response*, defined as the proportion of patients initially assigned to each condition who completed and responded to acute treatment and who then remained free from relapse across the first 12 months of follow-up, adjusted for censored observations (see Hollon et al., 2005). Although translating serial outcomes into a single dichotomous metric does sacrifice power to some extent, a dichotomous metric has the advantage of being representative of the full "intent-to-treat" sample and is therefore not susceptible to bias due to differential retention.

Figure 4 presents the sustained response in each treatment condition at the end of the 1st (full bar) and 2nd (lower portion of prior CT and BA bars) years of follow-up. Only 20% of the cPLA patients evidenced a sustained response through the end of the 1st follow-up year, compared to 23% of the cADM patients. By way of contrast, 44% of the patients initially assigned to prior CT and 34% of the patients initially assigned to prior BA experienced a sustained response through the end of the 1st follow-up year. What this means is that over one third of the patients initially assigned to psychotherapy showed sustained outcomes across the course of acute treatment and the 1st follow-up year, compared to less than a quarter of the patients initially assigned to pharmacotherapy. Including all patients initially assigned to treatment ($N = 188$),
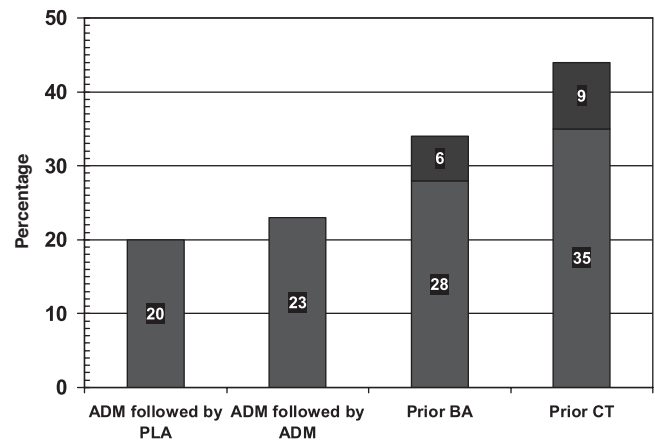

Figure 4. Sustained improvement for all patients initially assigned to active treatment. The total percentages of each bar represent the sustained response in each treatment condition at the end of the 1st year of follow-up; the lower portion of the Prior CT and Prior BA bars represents the sustained response (i.e., sustained recovery) at the end of the 2nd year of follow-up. ADM = antidepressant medication; PLA = placebo condition; BA = behavioral activation; CT = cognitive therapy.

sustained response varied significantly across groups, $\chi^2(3, N = 188) = 7.87$, $p < .05$. Pairwise comparisons revealed that only prior CT had a greater sustained response than both cADM, $\chi^2(1, N = 58) = 5.03$, $p = .02$, and cPLA, $\chi^2(1, N = 51) = 5.78$, $p = .016$. Pooling across the different types of treatment, the effects of the two prior psychotherapies were also significantly different than the effects in the medication conditions, $\chi^2(1, N = 106) = 6.69$, $p = .01$.

Across the 2nd follow-up year, rates of *sustained recovery* (recovered patients who were subsequently free of recurrence) were 35% for prior CT and 28% for prior BA. Although rates of sustained recovery were not calculated for the prior cADM patients, as they were withdrawn from medication at that point, the rate of sustained recovery could have been no better than the 23% rate of sustained response shown by this group at the start of the 2nd follow-up year, even if they had been maintained on medication. These results indicate that brief treatment with either CT or BA is as efficacious over the long run as keeping people on ADM.

Comparative Costs of Treatment

Figure 5 presents the direct costs of the psychotherapies compared to medication treatment. The cost estimates for this trial represent approximate costs in Seattle, WA at the time of the trial. Psychotherapy costs were about $2,000 per patient during the acute treatment phase (estimated at an average of 20 sessions per patient, at $100 per 50-min session). Over that same interval, ADM cost about $1,000 to provide per patient, with about half of that expense going to pay for physician time (12 sessions per patient at $75 per 20-min session) and the other half going to pay for medications (about $125 per month). However, from the end of acute treatment forward, the costs for ADM continued largely as before (session frequency was reduced to monthly sessions, but medication costs remained unchanged). As can be seen in the figure, the cumulative costs intersected at the 9th month of con-
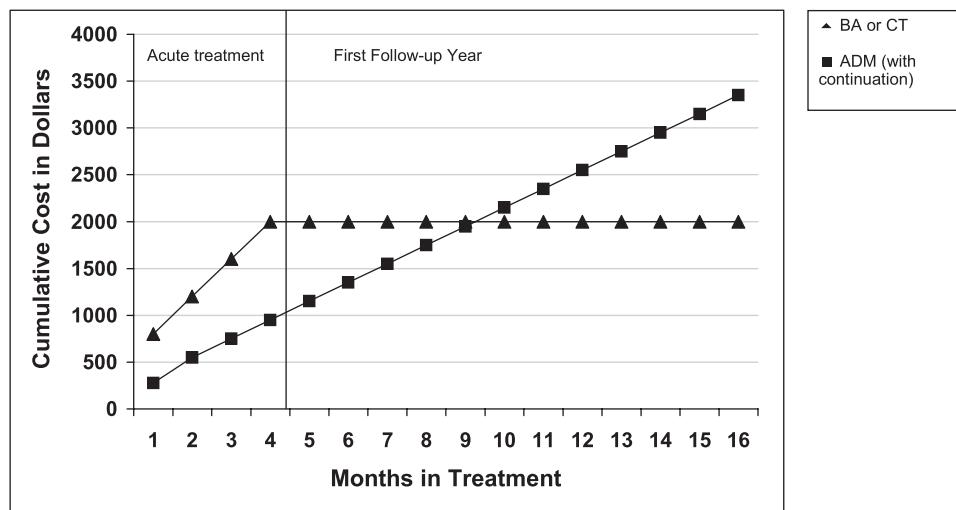
Figure 5. Cumulative direct costs of prior psychotherapy versus continued medication conditions. BA = behavioral activation; CT = cognitive therapy; ADM = antidepressant medication.

tinuing medication treatment, as ongoing ADM became more expensive than prior psychotherapy.

Power Considerations

Only 106 treatment responders were available to enter the 1st year of the follow-up, and only 48 recovered patients survived without relapse (or attrition) to complete the 2nd year. This loss to follow-up raised specific concerns about the power of the comparisons that involved the medication withdrawal condition (cPLA) during the 1st year of the follow-up, to which only 21 patients were assigned, and the BA condition during the 2nd year of follow-up, which consisted of only 13 recovered patients. On the basis of our observed sample sizes and recurrence rates within the psychotherapy conditions, we had 61.1% power to detect a clinically meaningful difference in recurrence rates, defined as a 30% difference between the prior psychotherapy and prior medication groups (Lakatos, 1988).

Discussion

The overall pattern of results observed in this study was that prior psychotherapy, either in the form of cognitive therapy (CT) or behavioral activation (BA), had an enduring effect that was at least as efficacious as continuing patients on medication and that held for the prevention of relapse and possibly recurrence. Evidence for this enduring effect was clearer for prior CT than for BA (differences relative to medication withdrawal were typically fully significant for the former, with nonsignificant trends for the latter), but differences between the two psychosocial interventions never approached statistical significance and were relatively small in magnitude. Although psychotherapy was more expensive to provide initially, and although no formal statistical comparisons of costs among treatment conditions were conducted, the cumulative cost of continued medications proved to be more expensive by the end of the 1st year of follow-up in this study. To the extent that the effects of psychotherapy endure over time, as is clearly the case for

CT and appears to be the case for BA, the cost savings could be considerable.

CT is one of the best-supported treatments for depression (De-Rubeis & Crits-Christoph, 1998) and is the only psychotherapy to date that has demonstrated an enduring effect in the treatment of depression (Hollon et al., 2006). For this reason, the indication that BA also may have an enduring effect comparable to CT, but not for patients successfully treated with medication, is particularly noteworthy. Moreover, these indications build on the findings from the acute phase of the trial, in which both BA and ADM each produced superior results to CT for more severely depressed patients (Dimidjian et al., 2006).

If BA does have an enduring effect, how might it achieve this result? BA emphasizes the role of behavior change in overcoming depression, consistent with the BA principles of increasing activity and approaching, rather than avoiding, difficult situations (Martell et al., 2001). Because these ideas are instantiated repeatedly during acute treatment, they are highly salient, and thus recall is increased at times of potential relapse. Approaching and solving problems during treatment also may fundamentally alter the environmental context in which the person lives, potentially reducing the likelihood of negative events that trigger risk for depression. In effect, BA is implemented in a manner that is intended to both teach coping skills and reduce future risk. The same is true for CT, which adds an emphasis on cognitive change, but otherwise takes a similar skills-training approach (Coffman, Martell, Dimidjian, Gallop, & Hollon, 2007). Both BA and CT had sustained responses that were not significantly different across the current trial. The mechanisms that underlie the enduring effect for CT are still not well understood (Hollon et al., 2006). Future research is needed to evaluate potentially common and unique mechanisms of change in each approach.

Medication treatment is considered the standard of care for depression in current psychiatric guidelines, with an indefinite length of treatment recommended for those with chronic or recurrent episodes (American Psychiatric Association, 2000; Frank et

al., 1990). However, although antidepressant medications generally are safe and efficacious, there is little evidence that they alter the course of the disorder. Even recovered patients are at substantial risk for recurrence once they stop taking medication (Hollon et al., 2002). Because depression is often chronic or recurrent (Kessler et al., 2003), any treatment that has an enduring effect is particularly worthwhile. In the current study, even though we found little evidence of a preventive effect for the continuation of medication, it was striking how rapidly even recovered patients experienced a recurrence when medication was withdrawn. In contrast, CT provided evidence of sustained benefit, and there were nonsignificant trends in the same direction for BA.

Beyond the enduring clinical benefit of CT and possibly BA, these findings suggest that there may be additional implications for healthcare delivery with regard to treatment costs. As suggested by Antonuccio et al. (1997), if psychotherapy truly has enduring effects, it may prove to be more cost effective than long-term medication treatment. Although more elaborate models for econometric evaluations of treatments exist (Ramsey et al., 2005), our relatively simple cost data suggest that the psychosocial interventions may be less expensive over the long run than keeping patients on medication. These findings also were consistent with direct cost estimates recently reported by Hollon et al. (2005). The current cost estimates were limited in time and place to when and where the study was conducted; however, the findings do suggest that more sophisticated econometric analyses should be considered in future trials.

There are several limitations that are worthy of comment. First, the study was conducted in the setting where BA was first developed, so it is possible that investigator allegiance may have biased the study in favor of that modality. We attempted to ensure that the other modalities were adequately implemented by involving investigators with allegiances to those approaches, and we appeared to succeed to the extent that we found strong acute effects for ADM and an enduring effect for CT, but replication of the study in other settings would be of significant value to the field. Second, any long-term follow-up study can be biased by differential retention among the treatment groups. Only slightly more than half of the patients originally randomized into active treatment during the acute phase of the trial were eligible for participation in the subsequent follow-up phase, which raises the possibility that differential retention may have biased the results (Klein, 1996). Although our analyses related to the issue of differential retention did not find any reason for concern, this possibility cannot be dismissed with respect to characteristics that we did not consider.

Third, although the pharmacotherapy conditions were conducted triple blind in the 1st year of follow-up, we did not assess the success of maintaining the blind. Fourth, we conducted more frequent assessments at the beginning of each follow-up year, when patients were being brought off medication, and it is possible that we may have missed later relapses or recurrences. Nonetheless, we made every effort to obtain a continuous record of symptom status over time, and patients in each condition were assessed according to the same common schedule. Fifth, there were also a number of constraints on the implementation of the treatments for research purposes that do not reflect usual patterns of clinical practice (cf. Schulberg et al., 1997). For instance, pharmacotherapists were limited to the use of a single medication throughout the trial, the taper schedule was relatively brief, and psychotherapists were required to terminate treatment at the end of the acute phase and not permitted to offer booster sessions or other ongoing contacts. Increased flexibility of the treatment protocols could be used in future trials to maximize the external validity of the findings.

Finally, aspects of the follow-up portion of the study lacked power for some comparisons. We had originally planned to run about half again as many subjects as we did, but the grant was not renewed in the wake of Neil S. Jacobson's untimely death, thus limiting the overall sample size. Given these constraints, it is notable that we were able to detect differences between the prior psychotherapies and medication conditions. Given the limited study power, these results speak both to the promise of BA and to the importance of replication.

Although the current results are promising, they also highlight areas of concern for future research. Although the two psychosocial interventions had enduring effects that were comparable to keeping patients on medication, slightly less than half of the patients initially randomized showed sustained clinical response in the best of treatments, and only about one quarter to one third of those patients evidenced sustained recovery across the full 2 years. Therapists in this trial were highly experienced and worked under conditions often considered ideal, including the provision of frequent supervision by experts. These benefits are rarely available in community settings (Westen, Novotny, & Thompson-Brenner, 2004). Thus, research for the continued improvement of psychosocial treatments for depression, particularly with a focus on recovery and prevention of relapse and in settings with higher ecological validity, is warranted (Dozois & Dobson, 2004).

Overall, the current results suggest that BA may have an enduring effect that approaches that produced by CT in the treatment of major depression. Prior CT was clearly superior to medication withdrawal, and prior BA did almost as well (at the level of a nonsignificant trend). Each was at least as effective as continued medication and at a lower cost. BA also was more efficacious than CT in the acute treatment of more severely depressed patients in this trial (Dimidjian et al., 2006). In aggregate, these results suggest that behavioral interventions for depression may deserve greater consideration than they have received in recent decades.

References

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

American Psychiatric Association. (2000). Practice guidelines for the treatment of patients with major depressive disorder (revision). *American Journal of Psychiatry, 157,* 1–45.

Antonuccio, D. O., Thomas, M., & Danton, W. G. (1997). A cost-effectiveness analysis of cognitive behavior therapy and fluoxetine (Prozac) in the treatment of depression. *Behavior Therapy, 28,* 187–210.

Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). *Cognitive therapy of depression.* New York: Guilford Press.

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the BDI-II.* San Antonio, TX: The Psychological Corporation.

Coffman, S., Martell, C. R., Dimidjian, S., Gallop, R., & Hollon, S. D. (2007). Extreme non-response in cognitive therapy: Can behavioral activation succeed where cognitive therapy fails? *Journal of Consulting and Clinical Psychology, 75,* 531–541.

Collett, D. (1994). *Modeling survival data in medical research.* New York: Chapman & Hall.

Cox, D. R., & Oakes, D. (1984). *Analysis of survival data.* London: Chapman & Hall.

DeRubeis, R. J., & Crits-Christoph, P. (1998). Empirically supported individual and group psychological treatments for adult mental disorders. *Journal of Consulting and Clinical Psychology, 66,* 37–52.

Dimidjian, S., Hollon, S. D., Dobson, K. S., Schmaling, K. B., Kohlenberg, R. J., Addis, M. E., et al. (2006). 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, 74,* 658–670.

Dozois, D. J. A., & Dobson, K. S. (Eds.). (2004). *Prevention of anxiety and depression.* Washington, DC: American Psychological Association.

Evans, M. D., Hollon, S. D., DeRubeis, R. J., Piasecki, J. M., Grove, W. M., Garvey, M. J., & Tuason, V. B. (1992). Differential relapse following cognitive therapy and pharmacotherapy for depression. *Archives of General Psychiatry, 49,* 802–808.

Fawcett, J., Epstein, P., Fiester, S. J., Elkin, I., & Autry, J. H. (1987). Clinical management: Imipramine/placebo administration manual. *Psychopharmacological Bulletin, 23,* 309–332.

Frank, E., Kupfer, D. J., Perel, J. M., Cornes, C., Jarrett, R. B., Mallinger, A. G., et al. (1990). Three-year outcomes for maintenance therapies in recurrent depression. *Archives of General Psychiatry, 47,* 1093–1099.

Frank, E., Prien, R. F., Jarrett, R. B., Keller, M. B., Kupfer, D. J., Lavori, P. W., et al. (1991). Conceptualization and rationale for consensus definitions of terms in major depressive disorder: Remission, recovery, relapse, and recurrence. *Archives of General Psychiatry, 48,* 851–855.

Gloaguen, V., Cottraux, J., Cucherat, M., & Blackburn, I. M. (1998). A meta-analysis of the effects of cognitive therapy in depressed patients. *Journal of Affective Disorders, 49,* 59–72.

Gortner, E. T., Gollan, J. K., Dobson, K. S., & Jacobson, N. S. (1998). Cognitive-behavioral treatment for depression: Relapse prevention. *Journal of Consulting and Clinical Psychology, 66,* 377–384.

Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery and Psychiatry, 23,* 56–61.

Hollon, S. D., DeRubeis, R. J., Shelton, R. C., Amsterdam, J. D., Salomon, R. M., O'Reardon, J. P., et al. (2005). Prevention of relapse following cognitive therapy vs. medications in moderate to severe depression. *Archives of General Psychiatry, 62,* 417–422.

Hollon, S. D., & Shelton, R. C. (2001). Treatment guidelines for major depressive disorder. *Behavior Therapy, 32,* 235–258.

Hollon, S. D., Stewart, M. O., & Strunk, D. (2006). Cognitive behavior therapy has enduring effects in the treatment of depression and anxiety. *Annual Review of Psychology, 57,* 285–315.

Hollon, S. D., Thase, M. E., & Markowitz, J. C. (2002). Treatment and prevention of depression. *Psychological Science in the Public Interest, 3,* 1–39.

Jacobson, N., Martell, C., & Dimidjian, S. (2001). Behavioral activation treatment for depression: Returning to contextual roots. *Clinical Psychology: Science and Practice, 8,* 255–270.

Jacobson, N. S., Dobson, K. S., Truax, P. A., Addis, M. E., Koerner, K., Gollan, J. K., et al. (1996). A component analysis of cognitive-behavioral treatment for depression. *Journal of Consulting and Clinical Psychology, 64,* 295–304.

Kaplan, E. L., & Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association, 53,* 457–481.

Keller, M. B., Lavori, P. W., Friedman, B., Nielsen, E., Endicott, J., McDonald-Scott, P., & Andreasen, N. C. (1987). The Longitudinal Interval Follow-up Evaluation. *Archives of General Psychiatry, 44,* 540–548.

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Koretz, D., Merikangas, K. R., et al. (2003). The epidemiology of major depressive disorder: Results from the National Comorbidity Survey Replication (NCS-R). *Journal of the American Medical Association, 289,* 3095–3105.

Klein, D. F. (1996). Preventing hung juries about therapy studies. *Journal of Consulting and Clinical Psychology, 64,* 81–87.

Kovacs, M., Rush, A. J., Beck, A. T., & Hollon, S. D. (1981). Depressed outpatients treated with cognitive therapy or pharmacotherapy: A one-year follow-up. *Archives of General Psychiatry, 38,* 33–39.

Lakatos, E. (1988). Sample sizes based on the log-rank statistic in complex clinical trials. *Biometrics, 44,* 229–241.

Martell, C. R., Addis, M. E., & Jacobson, N. S. (2001). *Depression in context: Strategies for guided action.* New York: Norton.

Nezu, A. M., Ronan, G. F., Meadows, E. A., & McClure, K. S. (2000). *Practitioner's guide to empirically based measures of depression.* New York: Kluwer.

Ramsey, S., Willke, R., Briggs, A., Brown, R., Buxton, M., Chawla, A., et al. (2005). Good research practices for cost-effectiveness analysis alongside clinical trials: The ISPOR RCT-CEA Task Force Report. *Value in Health, 8,* 521–533.

Schulberg, H. C., Block, M. R., Madonia, M. J., Scott, C. P., Lave, J. R., Rodriguez, E., et al. (1997). The "usual care" of major depression in primary care practice. *Archives of Family Medicine, 6,* 334–339.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86,* 420–428.

Simons, A. D., Murphy, G. E., Levine, J. L., & Wetzel, R. D. (1986). Cognitive therapy and pharmacotherapy for depression: Sustained improvement over one year. *Archives of General Psychiatry, 43,* 43–48.

van Schaik, D. J. F., Klein, A. F. J., van Hout, H. P. J., van Marwijk, H. W. J., Veekman, A. T. F., de Haan, M., et al. (2004). Patients' preference in the treatment of depressive disorder in primary care. *Psychiatry and Primary Care, 26,* 184–189.

Westen, D., Novotny, C. M., & Thompson-Brenner, H. (2004). Empirical status of empirically supported psychotherapies: Assumptions, findings, and reporting in controlled clinical trials. *Psychological Bulletin, 130,* 631–663.