Prevention of Relapse Following Cognitive Therapy vs Medications in Moderate to Severe Depression

Steven D. Hollon, PhD; Robert J. DeRubeis, PhD; Richard C. Shelton, MD; Jay D. Amsterdam, MD; Ronald M. Salomon, MD; John P. O’Reardon, MD; Margaret L. Lovett, MEd; Paula R. Young, PhD; Kirsten L. Haman, PhD; Brent B. Freeman, BA; Robert Gallop, PhD

Background: Antidepressant medication prevents the return of depressive symptoms, but only as long as treatment is continued.

Objectives: To determine whether cognitive therapy (CT) has an enduring effect and to compare this effect against the effect produced by continued antidepressant medication.

Design: Patients who responded to CT in a randomized controlled trial were withdrawn from treatment and compared during a 12-month period with medication respondents who had been randomly assigned to either continuation medication or placebo withdrawal. Patients who survived the continuation phase without relapse were withdrawn from all treatment and observed across a subsequent 12-month naturalistic follow-up.

Setting: Outpatient clinics at the University of Pennsylvania and Vanderbilt University.

Patients: A total of 104 patients responded to treatment (57.8% of those initially assigned) and were enrolled in the subsequent continuation phase; patients were initially selected to represent those with moderate to severe depression.

Interventions: Patients withdrawn from CT were allowed no more than 3 booster sessions during continuation; patients assigned to continuation medication were kept at full dosage levels.

Main Outcome Measures: Relapse was defined as a return, for at least 2 weeks, of symptoms sufficient to meet the criteria for major depression or Hamilton Depression Rating Scale scores of 14 or higher during the continuation phase. Recurrence was defined in a comparable fashion during the subsequent naturalistic follow-up.

Results: Patients withdrawn from CT were significantly less likely to relapse during continuation than patients withdrawn from medications (30.8% vs 76.2%; P=.004), and no more likely to relapse than patients who kept taking continuation medication (30.8% vs 47.2%; P=.20). There were also indications that the effect of CT extends to the prevention of recurrence.

Conclusions: Cognitive therapy has an enduring effect that extends beyond the end of treatment. It seems to be as effective as keeping patients on medication.

Arch Gen Psychiatry. 2005;62:417-422

Author Affiliations:
Departments of Psychology (Dr Hollon) and Psychiatry (Drs Shelton, Salomon, and Haman and Ms Lovett), Vanderbilt University, Nashville, Tenn; Departments of Psychology (Dr DeRubeis) and Psychiatry (Drs Amsterdam, O’Reardon, and Young and Mr Freeman), University of Pennsylvania, Philadelphia; and Department of Mathematics and Applied Statistics, West Chester University, West Chester, Pa (Dr Gallop).
during effect would be obtained with more severely depressed outpatients. Since the publication of findings from the Treatment of Depression Collaborative Research Program, questions have been raised about the effectiveness of CT with more severely depressed patients. To our knowledge, no other published trial has focused specifically on psychosocial treatment in this subpopulation. The present study asks whether CT has an enduring effect that extends to the prevention of relapse among more severely depressed outpatients, and it allows for a comparison of the magnitude of CT’s prevention effect relative to cADM.

METHODS

This study examines the subsequent course following initial treatment for patients randomized to either CT or ADM. A placebo-controlled continuation design was used to compare patients who responded to 16 weeks of CT with patients who responded to 16 weeks of ADM. Subjects were patients with moderate to severe unipolar depression aged 18 to 70 years who were recruited from outpatient psychiatric clinics at 2 sites, the University of Pennsylvania and Vanderbilt University. The full details of the screening process and the patient characteristics are given elsewhere. Institutional review boards at the University of Pennsylvania and Vanderbilt University reviewed and approved the study, including the withdrawal of active medication from patients who responded to treatment. Written informed consent was obtained from all participants, so that all of them knew that treatment might be withdrawn shortly after their initial response. All patients met the criteria for major depressive disorder as ascertained by the Structured Clinical Interview for DSM-IV-TR diagnoses. Moreover, they had to have scores of 20 or above for 2 consecutive weeks on the first 17 items of the Hamilton Depression Rating Scale (HDRS). This was the criterion used by Elkin and colleagues to define patients as a responder. Because all HDRS scores of 12 or less. These criteria prevented a transient 12-week HDRS scores of 12 or less; or (2) weeks 12, 14, and 18 were enrolled into the 12-month continuation phase of the study.

Of the 180 patients who had been assigned to one of the active treatments, 104 (57.8%) met the criteria for response and the remaining 60 patients received 8 weeks of pill placebo (cP-P) and treatment with either CT (n=60) or ADM (n=120); the re- mission criteria. They were randomly assigned to 16 weeks of acute treatment with either cADM (n=34) or withdrawal onto cP-P (n=35). They were monitored closely for the reemergence of depressive symptoms during this time, as were patients who had responded to CT (n=35). All patients were asked not to pursue treatment for depression other than that provided in the research protocol during the yearlong continuation phase. Patients who completed the continuation phase without relapse were withdrawn from all treatment and observed across a subsequent yearlong naturalistic follow-up.

TREATMENT

Clinical Management and Drug Continuation

Patients treated with ADM continued with the same psychiatrist they saw for acute treatment. Sessions were held at least every 2 weeks for the first month of continuation, and at least monthly thereafter. Clinical management sessions typically lasted about 15 to 30 minutes, and were conducted in accordance with the manual used in the Treatment of Depression Collaborative Research Program. Jan Fawcett, MD, the author of that manual, provided training in clinical management before the study began, and consultation on its implementation during the study. Session content focused on symptoms and adverse effects. Limited advice giving was allowed, and support was provided. Techniques and strategies specific to CT were prohibited.

Patients typically had been treated with paroxetine during acute treatment; treatment in those who had experienced less than a full response by 8 weeks was augmented with lithium or desipramine hydrochloride. Patients who were randomly assigned to stay on medications during the continuation phase typically continued on the same medications and dosages to which they responded, although dosage reduction was allowed as a means of dealing with adverse effects. In a few cases, medications were switched or augmented to deal with the reemergence of depressive symptoms.

Patients who were withdrawn onto cP-P continued to meet with their treating psychiatrist on the same schedule as previously described, and continued to receive placebos identical in appearance to the medications to which they had initially responded. Placebos were phase in during a 4- to 6-week period, with the dose of paroxetine typically reduced in 10-mg decrements weekly. Withdrawal onto cP-P was conducted on a blinded basis; patients, psychiatrists, and evaluators were all kept blind as to whether the patient was receiving an active medication or a placebo. Adjustments to medication doses for patients taking placebo were handled in the same manner as for patients taking active medications. Fabricated plasma levels were provided to the treating psychiatrists for those patients augmented with placebo lithium during the continuation phase, to maintain the blinding.

Booster Sessions (CT)

Although responders to CT discontinued treatment at the end of the acute phase, they were allowed up to 3 booster sessions...
cal symptoms.15 (An additional 7 items, including 3 that em-
judged to have relapsed if he or she was diagnosed as having
view of full major depressive disorder criteria. A patient was
the preceding interval; elevated scores are used to trigger a re-
for level of depression on a 6-point scale for each week during
Longitudinal Interval Follow-up Evaluation, patients are rated
nation yielded a
ment of the reliability of the major depressive episode desig-
ations were videotaped. Interviewers at both sites (4 at Uni-
weeks) at any time during the continuation period. All exami-
ments were conducted more frequently at the beginning of the
through the end of month 2, and monthly thereafter. Assess-
the domain of CT. Some sessions involved crisis intervention,
would be needed. Session content was left free to vary within
the start of the study. Nevertheless, the
sample considered herein was still marked by high lev-
order. Treatment responders also were more likely to be
sored because of a premature return to depression treat-
ment: 1 cADM patient insisted on adding psycho-
therapy because of functional impairment in month 2 and
CT patient began taking an ADM in month 10. Nei-
ther patient met the criteria for relapse either before or
after pursuing additional treatment, although the pa-
ient in the cADM group came close on several occa-
sions before and after her return to treatment.

**INITIAL CHARACTERISTICS OF TREATMENT GROUPS**

As a group, relative to those who did not enroll in the
continuation phase of this study, patients who com-
pleted and responded to treatment were less likely than
dropouts and nonresponders to have comorbid diag-
noses of posttraumatic stress disorder or cluster A per-
sonality disorders, especially paranoid personality dis-
order. Treatment responders also were more likely to be
employed at the start of the study. Nevertheless, the
sample considered herein was still marked by high lev-
als of comorbidity and chronic depression. More than 80%
of the sample met the criteria for at least 1 other disor-

**OUTCOME MEASURES**

The assessment schedule called for patients to be assessed weekly
by a blind clinical evaluator for the first 2 weeks, every other week
through the end of month 2, and monthly thereafter. Assess-
ments were conducted more frequently at the beginning of the
continuation phase to ensure that emerging relapses were not
missed among patients in whom treatment was being with-
drawn. In addition, patients were encouraged to call the clinic if
they were concerned that depressive symptoms were reemerg-
ing, in which case an ad hoc examination was scheduled as soon
as possible. Moreover, whenever a patient met the severity cri-
tion for relapse, an examination was scheduled for 1 week later,
to ascertain whether the temporal criterion was also met.

The primary measure used in the ascertainment of relapse
was a 17-item version of the HDRS,11 modified to include atypi-
cal symptoms.12 (An additional 7 items, including 3 that em-
phasize cognitive symptoms, were assessed, but not used to as-
certain relapse.) A patient met relapse criteria if he or she was
given a score of 14 or greater on the HDRS for 2 consecutive
weeks. In those instances when patients failed to come in for a
scheduled examination, or failed to notify project personnel
when they began to become symptomatic between reexamina-
tions, patients could also be judged to have met relapse crite-
ia based on the Longitudinal Interval Follow-up Evaluation,16
which was conducted at least every 3 months. On the
Longitudinal Interval Follow-up Evaluation, patients are rated
for level of depression on a 6-point scale for each week during
the preceding interval; elevated scores are used to trigger a re-
view of full major depressive disorder criteria. A patient was
judged to have relapsed if he or she was diagnosed as having
major depressive disorder (a score of ≥5 for 2 consecutive
weeks) at any time during the continuation period. All exami-
nations were videotaped. Interviewers at both sites (4 at Univer-
sity of Pennsylvania and 3 at Vanderbilt University) rated a
subset of these tapes. An intraclass correlation coefficient of 0.96
was obtained for the 17-item total HDRS score (n = 24). Assess-
ment of the reliability of the major depressive episode design-
ent yielded a k coefficient of 0.80 (n = 12).17

Once it was determined that a patient met the relapse cri-
tia, the onset of the relapse was dated to the point at which
those criteria were met, typically 2 weeks after symptom on-
set. Three weeks of increased symptoms were required to meet
the criteria for relapse during the first month of continuation,
so as not to misconstrue transient withdrawal symptoms as in-
dicative of a bona fide clinical relapse.

**DATA ANALYSIS**

To identify potential confounds in the relapse analyses, 5 demo-
ographic, 5 history of illness, 4 diagnostic subtype, 13 comor-
bidity, and 2 personality variables were examined to deter-
mine whether any of them differentiated the 3 follow-up
conditions and predicted relapse. By using a liberal P = .10, 2
indexes, dysthymia and atypical subtype, met these criteria.

Therefore, for the main relapse analyses, both were used as co-
variates. The number of prior episodes (which did predict sub-
sequent relapse) and sex (which did not predict subsequent re-
lapse) were also included as covariates, because both were used
as stratification variables during the randomization before con-
tinuation. Inclusion of the covariates did little to affect the com-
parisons involving pCT, but they did sharpen differences com-
paring cADM with pCT.

Survival curves and relapse rates were estimated using the
Cox proportional hazards regression model.18 As is typically
done, patients unavailable for follow-up were treated as cen-
sored observations, as were patients who returned to depres-
sion treatment without a documented relapse. Survival rates
for the 3 conditions were compared using the log-rank test. The
Cox proportional hazards regression model was used to evalu-
ate the influence of potential prognostic indexes, as described
by Collett.19 Statistical significance was set at P < .05 (2-tailed),
and specific contrasts were conducted among the respective
conditions using the Cox proportional hazards regression model
for the survival curves and Cox-Mantel-Haenszel analysis
controlling for site for relapse rates. In the presence of small
cell sizes, Fisher exact tests replaced the Cox-Mantel-
Haenszel analysis. Similar analyses were applied to data col-
lected during the naturalistic follow-up.

**RESULTS**

**PATIENT FLOW AND DROPOUT**

Sixteen patients dropped out of protocol during the con-	inuation phase, 8 from the cP-P group, 5 from the cADM

©2005 American Medical Association. All rights reserved.
As shown in Figure 1, prior exposure to CT reduced the risk for subsequent relapse relative to cP-P ($\chi^2 = 8.68, P = .01$). Relative to cP-P, cADM reduced relapse at the level of a nonsignificant trend ($\chi^2 = 3.14, P = .08$). Prior CT and cADM did not differ significantly ($\chi^2 = 1.62, P = .20$). Adjusted relapse rates for each condition were 30.8% for pCT, 47.2% for cADM, and 76.2% for cP-P. Hazard ratios were calculated between c-P and each of the respective active treatments. Prior exposure to CT was associated with a hazard ratio of 0.30 relative to c-P, which means that prior exposure to CT reduced risk by 70%. Continuation ADM was associated with a hazard ratio of 0.50 relative to c-P, which means that keeping patients on medications essentially cut risk by half. This is comparable to what has been reported elsewhere in the ADM continuation literature.

Four of the patients who relapsed in the cADM condition did so when they were not adhering to their medication; defined as taking $<75\%$ of the prescribed medication for at least 1 week during the month before relapse. Therefore, a second set of analyses was conducted in which these observations were censored for nonadherence. In these analyses, cADM significantly outperformed cP-P ($\chi^2 = 5.44, P = .02$). Taking nonadherence into consideration decreased the relapse rate for cADM to 42% and produced a hazard ratio of 0.37 relative to c-P. This denotes a reduction in risk of 63%, close to that produced by pCT.

**SUSTAINED RESPONSE**

We also examined the proportion of patients in each condition who showed sustained response, defined as completing and responding to acute treatment and staying free from relapse across the 12-month continuation phase, adjusted for censored observations. As shown in

**NATURALISTIC FOLLOW-UP**

A total of 40 patients who remained active in ongoing assessments completed the 12-month continuation phase without relapse. This included 20 patients in the pCT, 14 in the cADM, and 6 in the c-P group. These patients were observed for an additional year in a naturalistic follow-up; pCT patients were allowed no further booster sessions, and all pills were withdrawn from the patients in the cADM and c-P groups in accordance with the same schedule followed at the end of acute treatment. Given that these patients had gone 12 months without relapse following initial remission, they can be considered to have recovered from the index episode. Any subsequent return of symptoms would be considered a recurrence, the onset of a wholly new episode. In other respects, recurrence was defined in the same manner as relapse ($\geq 2$ weeks of increased symptoms on the HDRS or Longitudinal Interval Follow-up Evaluation). None of these patients were unavailable for follow-up. As shown in Figure 3, survival analyses indicated that CT’s enduring effect extended to the prevention of recurrence. In the pCT group, 5 of 20 patients had a recurrence during the naturalistic follow-up, vs 7 of the 14 cADM group patients withdrawn from medication; adjusted recurrence rates were 17.3% for the pCT vs 53.6% for prior cADM following withdrawal from medication ($\chi^2 = 6.81, P = .009$). The hazard ratio for this comparison was 0.15, meaning that prior exposure to CT reduced risk for recurrence by 85%. Although not depicted in the figure, 2 of 6 patients in whom c-P was withdrawn also experienced a recurrence. Although it would be inappropriate to extend analyses across the full 2-year follow-up in the absence of a maintenance medication condition, 15
(25.0%) of the 60 patients initially assigned to CT showed a sustained response free of either relapse or recurrence. Given that only 14 patients assigned to continuation medication ended continuation treatment free from relapse (23.3% of a total possible 60), cADM could have done no better than pCT even if patients who survived the continuation phase had been kept on maintenance medication.

The findings of this study suggest that CT has an enduring effect that reduces risk following successful treatment, as indicated by the reduced relapse rates relative to medication withdrawal. Moreover, the magnitude of the CT effect seems to be at least as great as that achieved by keeping patients on continuation medication, which is widely regarded as the most effective means of preventing relapse. Thus, it seems that there are at least 2 ways to protect patients against relapse following successful treatment: to either continue ADM or provide CT during acute treatment. Moreover, there are indications that the enduring effect of CT may extend to the prevention of recurrence.

These findings need to be interpreted cautiously. No one would recommend withdrawing ADM from depressed patients treated solely with medication after only 4 months of treatment. In this study, ADM was withdrawn and patients began to take cP-P solely to determine whether CT had an enduring effect.

To the extent that CT has an enduring effect, it might prove less costly than ADM to provide over time. Assuming costs of at least $100 per hour for 20 to 25 sessions of CT and $75 per hour for briefer pharmacotherapy sessions (and $125 per month for medications), CT costs about twice as much as ADM during a 4-month acute phase, but this gap is closed by the eighth month of continuation medication treatment, and is reversed beyond that point, such that direct treatment costs for ADM exceed those of CT thereafter. We did not make assessments of other direct or indirect costs that would have allowed us to conduct a sophisticated econometric analysis, but others who have compared CT with medications on such indexes have found that medications alone may result in a 33% higher expected cost than individual CT.22

It remains unclear just how CT exerts its enduring effect. Patients are trained from the start to “do the therapy for themselves” rather than to be passive recipients of the therapy. From the first session on, patients are encouraged to test the accuracy of their beliefs in homework assignments, and considerable time is devoted in later sessions to anticipating problems that are likely to arise after treatment is completed. Our impression is that patients initially need to apply the skills they learned during treatment in a concerted fashion, but that these compensatory strategies eventually become second nature, coinciding with a parallel change from problematic underlying beliefs to more adaptive ones. Such a change in beliefs would be expected to reduce the likelihood of becoming distressed in situations that formerly were problematic.23 This process might hold whether the actual mechanism was a change in the content of the beliefs or a change in the way that patients react to their thoughts.24

The present findings speak primarily to the prevention of relapse, the return of the treated episode. Although there were indications that CT’s enduring effect may extend to the prevention of recurrence, direct comparisons to maintenance medication in larger samples would be required to fully assess its relative value. Moreover, these findings also do not speak to the consequences of combining CT and ADM, although prior studies suggest that CT’s enduring effect is robust even when combined with medications. Both of these questions should be examined further.

Submitted for Publication: August 12, 2003; final revision received July 20, 2004; accepted September 9, 2004.

Correspondence: Steven D. Hollon, PhD, Department of Psychology, Vanderbilt University, 306 Wilson Hall, Nashville, TN 37203 (steven.d.hollon@vanderbilt.edu).

Funding/Support: This study was supported by grants MH55875 (R10) and MH01697 (K02) (Dr Hollon) and MH50129 (Dr DeRubeis) from the National Institute of Mental Health, Bethesda, Md.

Previous Presentation: This study was presented at the 155th Annual Convention of the American Psychiatric Association; May 23, 2002; Philadelphia, Pa.

Acknowledgment: We thank GlaxoSmithKline, Brentford, Middlesex, United Kingdom, for providing medications and pill placebos for the trial. We also thank the following colleagues for contributing to this research (the specific contributions of some authors are also noted). Drs Hollon and DeRubeis were the principal investigators and oversaw the implementation of CT at the respective sites. Drs Shelton and Amsterdam were the coprincipal investigators and supervised the implementation of medication treatment. Edward Schweizer, MD, provided consultation about study design and implementation, especially early in the trial. Ms Lovett and Dr Young served as the study coordinators. Drs Salomon and O’Reardon and the late Martin Szuba, MD, served as study pharmacotherapists (along with Drs Shelton and Amsterdam). Cory P. Newman, PhD, Karl N. Jamnach, PhD, Frances Shusman, PhD, and Sandra Seidel, MSN, served as the cognitive therapists (along with Drs Hollon and DeRubeis). Jan Fawcett, MD, provided consultation on
the implementation of clinical management pharmacotherapy. Aaron T. Beck, MD, Judith Beck, PhD, Christine Johnson, PhD, and Leslie Sokol, PhD, provided consultation on the implementation of CT. Madeline M. Gladis, PhD, and Dr. Homan oversaw the training of the clinical interviewers, and David Appelbaum, PsyD, Laurel L. Brown, PhD, Richard C. Carson, PhD, Barrie Franklin, PhD, Nana A. Landenberger, PhD, Jessica Londa-Jacobs, PhD, Julie L. Pickholz, PhD, Pamela Fawcett-Pressman, MEd, Sabine Schmid, MA, Ellen D. Stoddard, PhD, Michael Suminski, PhD, and Dorothy Tucker, PhD, served as project interviewers. Dr. Gallop and Andrew J. Tomarken, PhD, provided statistical consultation. Joyce Bell, BA, Mr. Freeman, Cara C. Grugan, BA, Nathaniel R. Herr, BA, Mary Hooper, BA, Miriam Hundert, Veni Li- nos, MSc, and Tynya Patton, MA, provided research support. Kelly Benis Vitousek, PhD, provided helpful comments on an earlier draft of the manuscript.

REFERENCES