Journal of Consulting and Clinical Psychology
2004, Vol. 72, No. 4, 625–635

0022-006X/04/$12.00 DOI: 10.1037/0022-006X.72.4.625

Effectiveness of Cognitive–Behavioral Treatment for Panic Disorder Versus Treatment as Usual in a Managed Care Setting

Michael E. Addis, Christina Hatgis, Aaron D. Krasnow, and Karen Jacob
Clark University

Leslie Bourne
Fallon Clinic

Abigail Mansfield
Clark University

Eighty clients enrolled in a managed care health plan who identified panic disorder as their primary presenting problem were randomly assigned to treatment by a therapist recently trained in a manual-based empirically supported psychotherapy (M. G. Craske, E. Meadows, & D. H. Barlow, 1994) or a therapist conducting treatment as usual (TAU). Participants in both conditions showed significant change from pre- to posttreatment on a number of measures. Those receiving panic control therapy (PCT) showed greater levels of change than those receiving TAU. Among treatment completers, an average of 42.9% of those in PCT and 18.8% in TAU achieved clinically significant change across measures. The results are discussed with reference to the dissemination of PCT and other evidence-based psychotherapies to clinical practice settings.

Over the last 3 decades, psychotherapy researchers have developed and tested a wide range of empirically supported psychosocial treatments (ESTs) for adult psychiatric disorders (Chambless & Hollon, 1998; DeRubeis & Crits-Christoph, 1998; Westen & Morrison, 2001). Although the results of outcome studies conducted in tightly controlled research settings have typically supported the efficacy of ESTs, few of these treatments have been evaluated in actual clinical practice settings. Several differences between the contexts of research and clinical practice raise questions about the generalizability of treatment effects demonstrated in tightly controlled clinical trials. Factors that may limit the use of ESTs in clinical practice include the acceptability of treatments to practitioners and consumers, costs of training, length of treatments, greater heterogeneity of clients in clinical practice versus research settings, and other hypothesized limits on the generalizability of treatment effects from controlled research to clinical practice (e.g., Addis & Krasnow, 2000; Fensterheim & Raw, 1996; Persons & Silberschatz, 1998; Silverman, 1996). Potential strengths of ESTs include actuarially based clinical decision-making guidelines, well-operationalized treatment interventions, and methods for evaluating practice outcomes (Addis, Wade, & Hatgis, 1999; Persons & Silberschatz, 1998; Schoenwald & Hoagwood, 2001; Wilson, 1996).

Unfortunately, the strength of opinions about the clinical utility of ESTs far surpasses the available empirical evidence; only a handful of studies have evaluated the effectiveness of ESTs in clinical practice settings.[1] Wade, Treat, and Stuart (1998) used a benchmarking strategy to examine the generalizability of a cognitive–behavioral treatment for panic disorder (Craske, Meadows, & Barlow, 1994) to a community mental health care setting and found outcomes comparable with previous controlled efficacy trials at posttreatment and at a 1-year follow-up (Stuart, Treat, & Wade, 2000). Similar results have been reported for cognitive–behavioral therapy (CBT) in the treatment of depression (Persons, Bostrom, & Bertagnolli, 1999), bulimia (Tuschen-Caffier, Pook, & Frank, 2001), agoraphobia (Hahlweg, Fiegenbaum, Frank, Schroeder, & von Witzleben, 2001), and obsessive–compulsive disorder (Franklin, Abramowitz, Kozak, Levitt, & Foa, 2000; Warren & Thomas, 2001). Fewer studies have included a treatment-as-usual (TAU) control group to evaluate the effectiveness of an EST compared with routine clinical practice. Morgenstern, Blanchard, Morgan, Labouvie, and Hayaki (2001) provided training in a cognitive–behavioral EST for substance abuse to practitioners working in a community setting. This study revealed no differences between a structured version of the EST, a flexibly implemented version, and a TAU control group on any outcome measure. In contrast, Turnkington, Kingdon, and Turner (2002) found that a brief CBT intervention for schizophrenia produced superior changes in depression, insight, and schizophrenic symptoms compared with TAU.

Michael E. Addis, Christina Hatgis, Aaron D. Krasnow, Karen Jacob, and Abigail Mansfield, Department of Psychology, Clark University; Leslie Bourne, Fallon Clinic, Worcester, Massachusetts.

Preparation of this article was supported by National Institute of Mental Health Grant R29 MH57778.

Correspondence concerning this article should be addressed to Michael E. Addis, Department of Psychology, Clark University, 950 Main Street, Worcester, MA 01610. E-mail: maddis@clarku.edu

[1] For the sake of brevity, we limit our review to effectiveness studies in mental health care specialty settings because these are the settings in which psychotherapy treatments are typically delivered. However, it should be noted that a large proportion of mental health services (e.g., medication, informal counseling, diagnosis and referral) are delivered in primary care settings that create different barriers and pathways to treatment dissemination.

Taken together, these studies suggest that the outcomes of cognitive–behavioral ESTs in clinical practice are comparable with those found in controlled efficacy studies. However, currently little is known about the effectiveness of ESTs compared with routine clinical care in different practice contexts. Studies that directly compare the outcomes of ESTs with existing clinical services provide one of the true litmus tests for the effectiveness of an empirically supported treatment (Addis, 1997; Addis et al., 1999; Chambless & Ollendick, 2000).

Our goal in the current study was to test the effectiveness of panic control therapy (PCT; Barlow & Craske, 1989; Craske et al., 1994), a well-established EST for panic disorder, compared with TAU in a capitated managed care practice context. PCT is an 8–15-session manual-based cognitive–behavioral treatment that integrates psychoeducational, cognitive, and exposure-based (agoraphobic and interoceptive) treatment components. The treatment begins with psychoeducation about the role of physiological, cognitive, and behavioral factors in the etiology and maintenance of panic disorder. Clients then learn to observe and challenge dysfunctional thinking (e.g., catastrophizing, overestimating the probability of dreaded events), to reduce overall levels of arousal by breathing diaphragmatically, and ultimately to gradually expose themselves to feared sensations and situations until habituation to feared stimuli can take place.

The efficacy of PCT and other similar cognitive–behavioral treatments for panic disorder has been well established in numerous randomized trials (Gould, Otto, & Pollack, 1995; White & Barlow, 2002). To our knowledge, only Wade et al.'s (1998) study has compared PCT outcomes in clinical practice with those found in efficacy studies, and no studies have tested the relative effectiveness of PCT compared with TAU in an actual clinical service setting. This is a critical issue because the potential of empirically supported psychotherapies to improve outcomes in clinical practice is perhaps the primary reason for their development, evaluation, and dissemination.

In the current study we compared the effectiveness of PCT and TAU in a capitated managed care context. An increasing proportion of mental health services throughout the country are delivered in managed care settings, and such settings create both pathways and obstacles to the dissemination of ESTs (Hatgis et al., 2001; Strosahl, 1998). For example, using ESTs is one strategy for enhancing cost-effectiveness and accountability in clinical practice. However, intensive productivity demands combined with session limits and other cost-containment strategies can negatively affect practitioner and consumer morale and create structural barriers to the implementation of ESTs. More generally, practitioners may find themselves in a culture of productivity, rather than in the culture of learning necessary to acquire and maintain clinical skills associated with empirically based practice. Nonetheless, given the superior track record of PCT over comparison treatments in efficacy studies, we predicted that clients receiving treatment from a therapist trained in PCT would show superior outcomes compared with those treated by a therapist conducting TAU.

Method

Participants

All participants were members of a large HMO serving a wide cross section of central Massachusetts. Recruitment took place from February 1999 through July 2000. Participants were either self-referred in response to notices in a member newsletter or referred by physicians or mental health professionals working within the HMO. Participants were eligible for the study if they met diagnostic criteria for panic disorder with or without agoraphobia or were subthreshold for a strict diagnosis of panic disorder but identified panic symptoms as their primary reason for seeking treatment. Participants were excluded if they were seeking treatment for a problem other than panic or anxiety, had an untreated substance-use problem in the last 6 months, had a diagnosis of psychosis in the past 5 years, were currently judged to be at risk for suicide, or were concurrently involved in other individual psychotherapy. No exclusions were made on the basis of medication use for anxiety or other comorbid psychological or medical problems.

One hundred fifty-five potential participants completed a diagnostic assessment with a trained interviewer (described below). Seventy-five were not enrolled in the study. Of these, 38 (51%) did not meet one of the study criteria. Two reported that panic disorder was not the primary problem for which they were seeking help, 2 reported that treatment would be too time-consuming, 1 objected to having sessions audiotaped, 5 were assigned as training cases to therapists recruited midstudy, and 29 did not provide a reason for choosing not to participate). The resulting sample consisted of 80 participants (70% were women and 30% were men, M age = 39.9 years, SD = 12.9, range = 18–70). Thirty-eight participants were randomly assigned to PCT and 42 to TAU.

Therapists

Shortly after the study started, and prior to therapist training, all outpatient mental health services provided by the HMO were carved out to an independent clinical service agency employing 24 master's-level therapists, all of whom were offered participation in the study. Master's-level practitioners were chosen because they provide a large and increasing percentage of outpatient mental health services in specialty clinics. Ten therapists agreed to participate. None of them identified their primary theoretical orientation as cognitive–behavioral; as a group they were approximately equally distributed between eclectic, family systems, psychodynamic, and humanistic in their self-described orientation. Of those who decided not to participate, the majority cited excessive caseloads, lack of available time for the 2-day PCT training workshop, or unwillingness to have sessions audiotaped. We created two groups of five therapists each matched on number of years in clinical practice and prior exposure to cognitive–behavioral treatment of anxiety or depression (e.g., attending a workshop). One group was randomly assigned (by the flip of a coin) to conduct PCT and the other to conduct TAU. Those in the TAU condition were told that they would receive training in PCT following completion of the acute treatment phase of the study. All therapists were informed of the processes for assigning treatment conditions (for both therapists and clients) prior to agreeing to participate in the study. Thus, all therapists knew that they would ultimately receive training in PCT.

Within the 1st year of recruitment, and following therapist training (described below), three therapists (two PCT, one TAU) stopped working at the clinic and subsequently left the study. Replacement therapists were drawn from the same population of clinic therapists, which included recent hires since the start of the study. None of the replacement therapists had previously participated in the study. Therapists were assigned to PCT or TAU on the basis of scheduling needs (e.g., case flow) in each condition and the need to maintain the balance of therapist characteristics between the groups. One doctoral-level therapist (PhD) was assigned to TAU. One doctoral-level therapist (PsyD) and one master's-level therapist (MA) were assigned to PCT. Those who joined the PCT condition reviewed videotapes of the initial workshop (described below). These therapists received the same session reviews and phone supervision as other therapists in the PCT condition. New PCT therapists were assigned two training cases from the ongoing flow of participant recruitment. Data from these cases were not included in subsequent outcome analyses. The therapist who joined the

TAU condition was given the same orientation and instructions as the initial group of therapists.

Training

No therapists in either condition received additional compensation beyond existing salaries or clinical fees for any time devoted to the study, including seeing cases, attending meetings, or receiving supervision during the training phase. Therapists in the PCT condition received a 2-day training workshop conducted by an expert with several years of experience in training clinicians from a variety of backgrounds in PCT techniques. In addition, PCT therapists were assigned two training cases. Each session was audiotaped, random sessions were reviewed, and therapists received four 30-min phone consultations with the expert who conducted the workshop. After consulting with several PCT researchers and trainers, we chose this level of training as a compromise between ensuring an adequate degree of supervision and creating a model of training that could be relatively easily disseminated to a variety of clinical practice contexts. In addition, PCT therapists attended twice-monthly 1-hr group consultation meetings with the principal investigator and research team. These meetings were used to discuss cases and refine therapists' knowledge of PCT and its underlying cognitive–behavioral principles. Because the study was designed as an effectiveness trial, mandating therapist attendance at consultation meetings would have been a serious threat to the generalizability of the findings. However, encouraging therapists to attend meetings by repeatedly expressing our confidence in the treatment and the value of consultation was consistent with maximizing the external validity of the study. Therapists chose what topics to discuss during each meeting and were not required to discuss specific cases as is more commonly the case in controlled efficacy studies. Nonetheless, the overwhelming majority of time was spent discussing therapeutic issues specific to PCT, such as how to modify interventions to meet the needs of particular clients, or clarifying the theory underlying the treatment as it applied to specific cases.

Treatments

PCT (Craske et al., 1994) is a manual-based 12–15-session cognitive–behavioral treatment protocol. The treatment includes education about the causes and maintenance of panic disorder, breathing retraining, cognitive restructuring, and interoceptive and agoraphobic exposure components. PCT therapists were encouraged but not required to use the treatment protocol. We explained to therapists that the treatment typically requires 12–15 sessions and is designed to be conducted weekly. However, therapists were not required to schedule a specific number or sequence of treatment sessions.

TAU therapists were instructed to provide whatever treatment they deemed appropriate for the clients they treated. In both the TAU and PCT conditions, we left decisions about medication use up to clients, their therapists, and other medical or psychiatric providers involved in their cases. Medication referrals during treatment generally were made to psychiatrists working within the clinic. However, some clients may also have received medication from primary care or other physicians.

Procedure

All participants provided written informed consent. Diagnostic interviewers explained the purpose of the study in a way designed to minimize any differences in the perceived credibility of the treatments. The following statement, taken directly from the consent form, describes the treatments as they were presented to participants:

> Which treatment you receive will be determined randomly, as if by flipping a coin. One of the therapies will be the standard treatment provided at (name of clinic). Standard treatment will be the best treatment for you as determined by you and your therapist at (name of clinic). Therapists vary in their approaches and your treatment may include supportive psychotherapy, cognitive–behavioral therapy, stress reduction techniques, exploring family and interpersonal issues, or other interventions your therapist thinks will be helpful. The other type of therapy available as part of this study will also be provided by therapists at (name of clinic). This therapy will involve cognitive–behavioral techniques for coping with anxiety, including education about panic and anxiety, and identification of thoughts and behaviors associated with panic. A more detailed description of this therapy will be provided by your therapist during your first session.

Clinic staff scheduled clients with a therapist in the randomly assigned treatment condition on the basis of availability and match between therapist and client schedules. Clients generally met with their therapist within 1 week of entrance into the study. Scheduling of the number and timing of subsequent therapy sessions was determined by therapist and client. Because treatment duration was left free to vary, we chose to assess participants after specified periods of time in treatment rather than after completion of a particular number of sessions. To time follow-up interviews, we assigned the date of 1 week after the initial assessment as the start of treatment. Participants were compensated $40 for a posttreatment assessment 5.5 months after treatment initiation. An additional 8.5-month follow-up was scheduled to serve as a brief assessment of panic symptoms 3 months after the posttreatment assessment. By the end of the 1st study year, it became apparent that approximately one third of the clients completed fewer than five sessions by the 5.5-month follow-up. Thus, analyses are presented for both posttreatment assessment periods.

Measures

Psychiatric diagnoses. Diagnostic and Statistical Manual of Mental Disorders (4th ed.; *DSM–IV*; American Psychiatric Association, 1994) diagnoses were based on the Anxiety Disorders Interview Schedule—IV (ADIS–IV; Brown, DiNardo, & Barlow, 1994). This structured interview is widely used in anxiety disorders treatment research, and previous versions have demonstrated good internal consistency and interrater reliability for panic disorder (DiNardo, Moras, Barlow, Rapee, & Brown, 1993). In addition to being used to diagnose the *DSM–IV* anxiety and mood disorders, the interview is used to screen for other disorders and problems such as psychosis, substance abuse, and psychological dysfunction due to medical conditions. Diagnostic interviews were conducted by advanced clinical psychology graduate students. Interviewers received extensive training and supervision in the use and scoring of the ADIS–IV from the principal investigator and from standardized training tapes. Diagnostic supervision took place in weekly meetings with the principal investigator in which the research team reviewed and, if needed, arrived jointly at diagnostic decisions.

Panic severity. The Panic Disorder Severity Scale (PDSS; Shear & Maser, 1994; Shear et al., 2001) is a 7-item semistructured interview that assesses overall severity of panic disorder and agoraphobic avoidance. Information is also gathered about severity of panic attacks, anticipatory anxiety, avoidance behavior, and work and social dysfunction. Psychometric evaluations suggest that the PDSS has good interrater reliability, moderate internal consistency, and good construct validity (Shear & Maser, 1994; Shear et al., 2001). As with the ADIS–IV, interviewers received weekly training and supervision on the scoring of the PDSS. This measure was completed at pretreatment, 5.5 months, and 8.5 months.

Phobic avoidance. The Fear Questionnaire (FQ; Marks & Matthews, 1978) is a 24-item self-report questionnaire assessing phobic avoidance on three anxiety subscales (agoraphobia, blood–injury, social phobia) and one general anxiety and depression subscale. For all analyses we used the total phobia score that combines items from the three anxiety subscales. Psychometric properties for the FQ have been reported as adequate (Cox, Parker, & Swinson, 1996), and the measure is recommended for studies assessing phobic avoidance in panic disorder (Cox, Swinson, Parker, Kuch, & Reichman, 1993; Shear & Maser, 1994). The FQ was completed at pretreatment and at the 5.5-month follow-up.

Depression. The Beck Depression Inventory (BDI–1; Beck, Ward, Mendelson, Mock, & Erbaugh, 1961) is a 21-item self-report questionnaire assessing severity of depressive symptoms. This measure is widely used, is highly correlated with other measures of depression such as the Hamilton Rating Scale (Hamilton, 1967), and has excellent psychometric properties (Beck, Steer, & Garbin, 1988). The BDI–1 was completed at pretreatment and at the 5.5-month follow-up.

General well-being. The Outcome Questionnaire (OQ–45; Lambert et al., 1996) is a 45-item self-report measure of symptom distress, social role functioning, and interpersonal functioning. We chose the OQ–45 as a measure of general well-being to complement the other more panic-specific measures. The measure demonstrates good internal consistency, test–retest reliability, and concurrent validity with other measures of psychiatric symptoms and interpersonal functioning (Lambert et al., 1996; Mueller, Lambert, & Burlingame, 1998; Umphress, Lambert, Smart, Barlow, & Clouse, 1997). We used the total score combining the three subscales in all analyses. Participants completed the measure at pretreatment and at the 5.5-month follow-up.

Treatment credibility. The treatment credibility questionnaire (TCQ) was derived from Borkovec and Nau's (1972) measure of the credibility of analogue therapy rationales. Participants completed the TCQ by rating the following six questions on a 7-point Likert-type scale: "How much do you expect this therapy to be helpful to you?" "How logical does your therapy seem to you?" "How scientific does your therapy seem to you?" "How much do you think your therapy will help you to understand the causes of your problems?" "How much do you think your therapy will help you to learn effective ways of coping with your problems?" and "To what degree does the treatment you are receiving match your ideas of what helps people in psychotherapy?" The scale yielded an internal consistency coefficient (α) of .86 in the current sample. Individual items were summed to create a single score, with higher scores indicating greater perceptions of treatment credibility. Participants completed the TCQ following their second therapy session, and therapists did not have access to clients' responses.

Demographic and medication information. We collected a variety of self-reported demographic information during the diagnostic evaluations including age, income, education, ethnicity, and previous treatment experience. Diagnostic evaluators documented participants' current use of medications for problems related to anxiety or depression. Because many medications (e.g., selective serotonin reuptake inhibitors) are used to treat both anxiety and mood disorders, for some participants it was difficult to determine the specific purpose of the medication. In these cases evaluators used clinical judgment based on the participant's self-report to determine the primary problem for which medication was being taken.

Therapist adherence. All therapy sessions were audiotaped, and a random sample in both conditions was rated to determine the degree to which the treatments differed in the frequency and intensity of PCT interventions. A rating manual was developed with detailed descriptions of 22 interventions, approximately half of which were specific to PCT. The remaining half were general cognitive–behavioral interventions not specific to PCT (e.g., assigning and reviewing homework) and other common therapeutic interventions (e.g., focusing on interpersonal relationships). Interventions were rated on a 5-point Likert-type scale combining two dimensions: (a) the amount of time spent using an intervention and (b) how closely the intervention resembled its description in the rating manual. Interrater reliabilities for the adherence measure are presented in the Results section.

Results

Sample Characteristics

Participants reported a mean age of 39.9 years (SD = 12.9, range = 18–70). Eighty percent identified themselves as Caucasian, 4% as African American, 2% as Hispanic, and 14% as other. Participants reported their annual family income to the nearest $1,000. The sample as a whole had a median income of $40,000. Twenty-five percent of participants received a high school diploma or passed a General Educational Development test, 20% completed some college courses, and 28% received a 2- or 4-year college degree.

Seventy-three percent of participants met criteria for panic disorder with agoraphobia, and 27% met criteria for panic disorder without agoraphobia. Twenty-one percent met criteria for one additional anxiety disorder diagnosis, 14% met criteria for two, and 9% for three. Thirty-nine percent of the sample also met criteria for current major depression. Sixty-five percent of participants reported currently using psychotropic medications at the pretreatment assessment.

Table 1 shows pretreatment characteristics for clients in the current study compared with those in Barlow, Gorman, Shear, and Woods's (2000) large, multisite efficacy study and Wade et al.'s (1998) effectiveness study comparing PCT in a community mental health setting with outcomes from previous controlled trials. Mean

Table 1
Pretreatment Characteristics for Participants in the Current Study Compared With Barlow et al. (2000) and Wade et al. (1998)

	Current study		Barlow et al. (2000)		Wade et al. (1998)[a]	
Pretreatment measure	*M*	*SD*	*M*	*SD*	*M*	*SD*
PDSS mean item score	1.84	0.84	1.83	0.54		
FQ-AG	12.5	9.0			16.5	10.8
FQ-SP	13.9	8.8			17.3	8.9
BDI–1	14.3	8.7			15.5	8.8
% using anxiolytic medication	56.3		0		60.0	
% using depression medication	37.5		0		20.0	

Note. PDSS = Panic Disorder Severity Scale; FQ-AG = Fear Questionnaire Agoraphobia subscale; FQ-SP = Fear Questionnaire Social Phobia subscale; BDI–1 = Beck Depression Inventory.
[a] These data are from "Transporting an Empirically Supported Treatment for Panic Disorder to a Service Clinic Setting: A Benchmarking Strategy," by W. A. Wade, T. A. Treat, and G. L. Stuart, 1998, *Journal of Consulting and Clinical Psychology, 66,* p. 236. Copyright 1998 by the American Psychological Association. Reprinted with permission of the author.

pretreatment PDSS scores in our sample were almost identical to clients in the Barlow et al. study. Clients in the current study showed somewhat lower levels of agoraphobic and specific phobic symptoms compared with the community sample in Wade et al.'s study. Mean pretreatment BDI–1 scores were comparable in the current study and Wade et al.'s study. The percentage of clients using anxiolytic medication was comparable in the current study (56.3%) and Wade et al.'s study (60.0%). Overall, the comparisons suggest that clients in the current study were roughly comparable with those treated in previous efficacy and effectiveness research.

Between-Treatments Therapist Comparisons

One-way analyses of variance (ANOVAs) revealed no differences between PCT and TAU therapists on age (PCT M = 40.2, SD = 5.3; TAU M = 36.0, SD = 10.5), $F(1, 11) < 1$, *ns*, or years in clinical practice (PCT M = 4.6, SD = 3.0; TAU M = 5.8, SD = 6.1), $F(1, 11) < 1$, *ns*. Prior to the start of the study, 14% of therapists in PCT (n = 1) and 33% in TAU (n = 2) had attended a workshop in CBT for anxiety (Fisher's exact test, *ns*). Fifty-seven percent (n = 4) in PCT and 50% (n = 3) in TAU had attended a workshop in CBT for depression (Fisher's exact test, *ns*). Differences in orientation were too varied relative to the number of therapists to allow for meaningful comparisons between the treatment conditions. However, as mentioned above, none of the therapists described their primary orientation as cognitive–behavioral.

Data Analytic Strategy

We began by testing for comparability between the two treatment conditions on a number of pretreatment variables. Continuous variables were analyzed with one-way ANOVAs, and Fisher's exact test was used for categorical measures. We used two complementary approaches for the primary outcome analyses. First, we conducted a series of mixed-design ANOVAs. PDSS, FQ, BDI–1, and OQ–45 scores served as dependent measures in four separate analyses. Treatment (PCT vs. TAU) served as a between-subjects factor, and time (pre- vs. posttreatment scores) served as a within-subjects factor. The specific effect of interest in these analyses was the Treatment × Time interaction term. A significant interaction would indicate that the degree of change in pre- versus posttreatment scores for a dependent measure varied according to treatment condition. Second, we used multivariate analyses of covariance (MANCOVAs) to test for differences between the treatment conditions at posttreatment (as opposed to change from pre to post as in the repeated measures analyses). MANCOVA was considered appropriate because pretreatment measures were completed before random assignment to treatment condition. Thus, there was little chance that potential covariates could be confounded with substantive aspects of the treatments (Miller & Chapman, 2001; Overall & Woodward, 1977). Pretreatment scores on specific measures were considered potential covariates if the effect size for treatment condition at pretest was greater than 0.20 (Cohen, 1977). Effect sizes for treatment condition are presented in Table 2. Homogeneity of regression tests indicated no differences between the treatment conditions in the regressions of dependent measures on potential covariates, further warranting the use of MANCOVA. Clinical significance analyses were also computed for the PDSS, OQ–45, FQ, and BDI–1 in each treatment condition.

Treatment Attendance Subsamples

Outcome analyses were conducted for three subsamples of participants. The full intent-to-treat sample included all participants enrolled into the study (N = 80) regardless of how many sessions they attended. Eleven participants (5 PCT, 6 TAU) did not attend any treatment sessions. All of these participants completed posttreatment assessments at 5.5 or 8.5 months, and their data were included in the intent-to-treat analyses. We conducted analyses on the intent-to-treat sample excluding those cases in which participants did not attend treatment. The results closely paralleled the full intent-to-treat sample. Results from the full sample are reported below. Primary outcome analyses were repeated for participants who attended between one and seven sessions by the 8.5-month follow-up (n = 37: 13 PCT, 24 TAU, Fisher's exact test, p = .05). We designated this sample treatment noncompleters, rather than dropouts, because therapy in TAU presumably was of varying duration depending on the clients' needs and the therapists' approaches. Finally, analyses were completed for a subsample of treatment completers who attended at least eight sessions by the 8.5-month follow-up[2] (n = 32: 20 PCT, 12 TAU, Fisher's exact test, p = .04). None of the training cases for therapists in the PCT condition were included in any of the subsequent analyses.

Pretreatment Comparisons, Treatment Credibility, and Attrition

Comparisons between treatment conditions on pretreatment measures for the intent-to-treat sample were conducted with a series of one-way ANOVAs. Mean pretreatment scores for each primary outcome measure in each condition are presented in Table 2. We found no significant differences between PCT and TAU on any of the pretreatment measures of anxiety, depression, or general well-being. The treatments did not differ on percentages of participants with comorbid axis I diagnoses or medication use at pretreatment. Perceptions of treatment credibility following the second session, as measured by the TCQ, also did not differ between the treatments (PCT M = 35.5, SD = 4.2; TAU M = 35.3, SD = 4.1).

At the 5.5-month follow-up, participants reported whether they judged themselves to have discontinued treatment prior to completing it. We considered participants who responded yes to this question and who completed at least one but fewer than six sessions to be treatment dropouts.[3] Fifteen participants (18.8%, 10 PCT, 5 TAU, Fisher's exact test, *ns*) fell into this category. Five participants expressed concerns about the treatment or rapport with their therapist, 5 had scheduling or time conflicts, 3 cited the

[2] The designation of TAU clients who attended at least eight sessions as completers is somewhat ambiguous because treatment in this condition was not expected to have any specific duration. Use of the term *completer* was driven by the need to equate subsamples of the two groups on treatment duration for analytic purposes.

[3] This definition of attrition deviates somewhat from efficacy studies in which number of sessions is the primary criterion for determining attrition. However, in clinical settings a client's subjective assessment of whether he or she appropriately or inappropriately terminated treatment plays an important role in considering whether someone has dropped out of treatment.

Table 2
Primary Outcome Analyses Comparing PCT With TAU on Anxiety, Depression, and General Well-Being in Three Participant Subsamples

	PCT			TAU			Time		Treatment × Time
Measure	*n*	*M*	*SD*	*n*	*M*	*SD*	*d*[a]	*F*(*df*), *p*	*F*(*df*), *p*
Intent-to-treat sample (*n* = 80)									
PDSS									
Pre	36	13.5	6.0	42	12.3	5.7	.20		
Post (5.5 months)	38	8.4	5.8	42	8.3	6.4	.02	$F(1, 78) = 54.1$***	$F(1, 78) < 1$, *ns*
Post (8.5 months)	38	5.3[b]	5.5	41	6.5[b]	5.8	.21	$F(1, 77) = 138.3$***	$F(1, 77) = 4.6$*
FQ									
Pre	37	39.4	24.2	42	38.1	25.1	.05		
Post	38	28.5	19.2	39	33.1	22.1	.22	$F(1, 74) = 12.1$**	$F(1, 74) = 2.2$, $p = .14$
BDI–1									
Pre	36	15.5	7.3	42	13.2	9.7	.25		
Post	38	10.0	8.1[b]	37	9.1	8.1	.12	$F(1, 71) = 27.2$***	$F(1, 71) = 4.3$*
OQ–45									
Pre	36	73.8	18.5	39	67.9	25.1	.27		
Post	35	60.4	23.3[b]	39	63.2	25.3	.12	$F(1, 68) = 9.9$**	$F(1, 68) = 5.2$*
Treatment noncompleters (*n* = 37)									
PDSS									
Pre	13	14.4	8.1	24	11.8	5.8	.39		
Post (5.5 months)	13	10.4	5.2	24	6.0	6.0	.72	$F(1, 35) = 30.4$***	$F(1, 35) = 1.05$, *ns*
Post (8.5 months)	13	6.2	6.4	23	5.2	5.6	.17	$F(1, 34) = 38.5$***	$F(1, 35) < 1$, *ns*
FQ									
Pre	13	46.5	23.4	24	37.0	28.9	.35		
Post	13	41.5	18.1	23	30.3	23.7	.50	$F(1, 34) = 3.7$, $p = .06$	$F(1, 34) < 1$, *ns*
BDI–1									
Pre	12	17.2	6.8	24	13.5	9.3	.43		
Post	12	9.9	5.0	22	9.3	8.9	.16	$F(1, 32) = 19.6$***	$F(1, 34) = 2.9$, *ns*
OQ–45									
Pre	12	79.2	19.0	21	68.2	25.2	.47		
Post	11	60.7	13.2	20	60.6	27.4	.08	$F(1, 29) = 10.4$**	$F(1, 29) = 2.1$, *ns*
Treatment completers (*n* = 32)									
PDSS									
Pre	20	12.9	5.0	12	12.0	5.7	.17		
Post (5.5 months)	20	7.0	5.8[b]	12	10.9	6.5	.62	$F(1, 30) = 12.2$**	$F(1, 30) = 5.8$*
Post (8.5 months)	20	4.0	4.3[b]	12	7.7	6.6[b]	.67	$F(1, 30) = 88.7$***	$F(1, 30) = 10.7$**
FQ									
Pre	20	35.5	25.2	12	38.7	21.0	.14		
Post	20	20.9	15.6[b]	11	36.4	17.3	.88	$F(1, 29) = 5.8$*	$F(1, 29) = 5.7$*
BDI–1									
Pre	20	14.1	7.4	12	11.0	11.2	.35		
Post	20	9.8	7.3	11	7.5	7.0	.32	$F(1, 29) = 4.3$*	$F(1, 29) = 1.1$, *ns*
OQ–45									
Pre	20	69.0	17.6	12	61.4	27.6	.35		
Post	20	58.5	26.4	11	59.2	20.5	.03	$F(1, 29) < 1$, *ns*	$F(1, 29) = 4.0$, $p = .06$

Note. PCT = panic control therapy; TAU = treatment as usual; PDSS = Panic Disorder Severity Scale; FQ = Fear Questionnaire Total Phobia Scale; BDI–1 = Beck Depression Inventory; OQ–45 = Outcome Questionnaire; pre = pretreatment; post = posttreatment.
[a] *d* refers to between-treatments effect sizes calculated according to Cohen (1977). [b] Significant change pre–post within a treatment condition.
* $p < .05$. ** $p < .01$. *** $p < .001$.

severity of their agoraphobia or anxiety as a factor, and 2 did not provide a specific reason for discontinuing treatment.

Primary Outcome Analyses

To determine possible covariates for the repeated measures analyses we correlated age, gender, number of comorbid diagnoses, and pretreatment medication use with pre–post change scores on the PDSS, FQ, OQ–45, and BDI–1. None of these correlations were statistically significant, and no covariates were included in the analyses.

Cases with missing data were deleted listwise for each analysis. Within the intent-to-treat sample, 1 case was deleted from the 8.5-month PDSS analysis, 4 cases from the FQ, 7 cases from the BDI–1, and 10 cases from the OQ–45. Within the treatment noncompleters, 1 case was deleted from the 8.5-month PDSS analyses, 1 case from the FQ, 2 from the BDI–1, and 6 from the OQ–45. Within the treatment completer sample, 1 case was de-

leted from the FQ analyses, 1 from the BDI–1, and 1 from the OQ–45. Cases with missing data did not differ from nonmissing cases on any of the primary outcome measures for which data were available.

Table 2 shows the results of a series of repeated measures ANOVAs for each of the three samples. For each analysis, time (pre–post) served as a repeated measures factor and treatment condition (PCT vs. TAU) as a between-subjects factor. When a significant Treatment × Time interaction was found, we conducted follow-up repeated measures analyses within each treatment condition to test whether significant change occurred in both groups on a particular measure.

In the intent-to-treat sample we found a significant effect for time on 5.5-month PDSS scores, suggesting change across both conditions from pre- to posttreatment, $F(1, 78) = 54.1, p < .0005$. A significant Treatment × Time interaction emerged for 8.5-month PDSS scores, $F(1, 77) = 4.6, p = .04$. Follow-up within-treatment analyses revealed significant change from pre- to posttreatment in both conditions. Inspection of the pre- and posttreatment means suggested that the degree of change was greater in PCT than in TAU. A significant interaction also emerged for BDI–1 scores, $F(1, 71) = 4.3, p = .04$, with follow-up analyses revealing significant pre–post change in PCT but not in TAU. Similar results were found for OQ–45 scores with a significant Treatment × Time interaction, $F(1, 68) = 5.2, p = .03$, and significant change in PCT but not in TAU. To test for absolute differences between the treatments at posttest (as opposed to change), we conducted a MANCOVA with posttreatment PDSS, FQ, BDI–1, and OQ–45 scores as dependent variables. Twelve cases were excluded because of missing data. The treatment condition effect was significant, $F(5, 59) = 2.5, p = .04$, with follow-up univariate analyses revealing significantly lower posttreatment scores in PCT on the FQ, $F(1, 63) = 8.0, p = .006$, and the 8.5-month PDSS, $F(1, 63) = 6.9, p = .01$.

Among treatment noncompleters, we found significant pre–post changes on all outcome measures with the exception of the FQ. There were no significant Treatment × Time interactions. We conducted a MANCOVA following the procedures describe above with seven cases excluded listwise because of missing data. This analysis yielded a nonsignificant treatment condition effect, $F(5, 20) = 1.3, p = .30$.

In the subsample of treatment completers, we found significant Treatment × Time interactions for the PDSS at 5.5 months, $F(1, 30) = 5.8, p = .02$, and 8.5 months, $F(1, 30) = 10.7, p = .003$. Within-treatment follow-up analyses revealed significant change in PCT but not in TAU on the 5.5-month scores. Both treatments produced significant change on 8.5-month scores with the Treatment × Time interaction suggesting greater change in PCT than in TAU. A significant Treatment × Time interaction also emerged for posttest FQ scores, $F(1, 29) = 5.7, p = .02$, with follow-up analyses revealing significant pre–post change in PCT but not in TAU. A MANCOVA with one case excluded listwise because of missing data revealed a significant treatment condition effect, $F(5, 23) = 3.4, p = .02$. Follow-up univariate tests indicated significantly lower posttreatment scores in PCT on the FQ, $F(1, 27) = 14.9, p = .001$.

Clinical Significance Analyses

We followed the methods described by Jacobson and Truax (1991) to determine the percentage of clients in each treatment condition who achieved reliable change and end-state functioning within a nonclinical distribution. Cut scores from published norms were obtained for the PDSS (Shear et al., 2001), the OQ–45 (Lambert et al., 1996), the FQ (Gillis, Haaga, & Ford, 1995), and the BDI–1 (Seggar, Lambert, & Hansen, 2002). Reliable change scores were calculated on the basis of the internal consistency or test–retest reliability of each measure. Table 3 presents the results of these analyses for each of the three subsamples. In the intent-to-treat sample, the average percentage (across all measures) of clients achieving clinically significant change was 31.7% in PCT and 22.0% in TAU. Among treatment completers, an average of 42.9% of clients in PCT and 18.8% in TAU achieved clinically significant change. For treatment noncompleters, the percentages were 20.9% for PCT and 28.3% for TAU.

Therapist Adherence to PCT

We rated therapist adherence for 67 of the 80 cases in the study. Data were missing for 11 cases in which the clients did not attend any treatment sessions and 2 cases in which the therapists had audiotaping difficulties. Cases with missing adherence data did not differ from the rest of the sample on any of the primary outcome measures at pre- or posttreatment.

Two advanced clinical graduate students and two postbaccalaureate research assistants were divided into two rating teams, and each team rated approximately 125 therapy sessions. All raters were blind to treatment condition. Thirty-nine sessions were rated by both teams, and the 1st session was rated for all clients. One

Table 3
Percentage of Clients Demonstrating Clinically Significant Change

	PCT		TAU		
Measure	*n*	%	*n*	%	Fisher's exact test
Intent-to-treat sample ($N = 80$)					
PDSS 5.5	11	28.9	13	31.0	*ns*
PDSS 8.5	22	57.9	20	48.8	*ns*
FQ	7	18.4	2	4.8	$p = .08$
BDI–1	7	18.4	4	9.5	*ns*
OQ–45	12	35.3	6	15.8	$p = .10$
Treatment noncompleters ($N = 37$)					
PDSS 5.5	1	7.7	11	45.8	$p = .02$
PDSS 8.5	6	41.2	13	56.5	*ns*
FQ	1	7.7	1	4.2	*ns*
BDI–1	3	23.1	3	12.5	*ns*
OQ–45	3	25.0	5	22.7	*ns*
Treatment completers ($N = 32$)					
PDSS 5.5	8	40.0	1	8.3	$p = .10$
PDSS 8.5	14	70.0	5	41.7	$p = .15$
FQ	5	25.0	0	0	$p = .13$
BDI–1	3	15.0	1	8.3	*ns*
OQ–45	20	64.5	11	35.5	$p = .11$

Note. PCT = panic control therapy; TAU = treatment as usual; PDSS = Panic Disorder Severity Scale, 5.5- or 8.5-month follow-up; FQ = Fear Questionnaire Total Phobia Scale; BDI–1 = Beck Depression Inventory; OQ–45 = Outcome Questionnaire.

session was randomly selected from Sessions 2 to 4, 1 from Sessions 5 to 8, 1 from Sessions 8 to 10, and 1 from any sessions beyond 10. Individual raters in each team separately listened to tapes, and teams met weekly to assess reliability and determine a consensus rating for each intervention. All four raters met biweekly to discuss any discrepancies in ratings and to assess reliability between the pairs. Intraclass correlations (ICCs) indicated good to excellent interrater reliability for each intervention in each team (Team 1 median ICC = .83, Team 2 median ICC = .74) and for reliability between teams (median ICC = .66).

We first computed total scores across treatment sessions for each intervention, and then computed mean total scores across interventions in three categories: interventions specific to PCT, general cognitive–behavioral interventions (e.g., general relaxation, changing all-or-none thinking), and other therapeutic interventions. We defined general cognitive–behavioral interventions as those consistent with CBT but not specific to PCT, such as challenging depressogenic automatic thoughts and beliefs, or self-monitoring of negative thoughts more generally. Specific intervention scores and mean category scores are presented by treatment condition in Table 4. PCT interventions were more common in PCT than in TAU, $F(1, 65) = 84.1$, $p < .0005$. PCT therapists scored higher than TAU therapists on all of the PCT interventions except for agoraphobic exposure, in which the frequency of use was low with no differences between the treatments. Therapists in PCT appeared to use more general CBT interventions than did TAU therapists, $F(1, 65) = 3.9$, $p = .05$. Within this category, PCT therapists scored higher on assigning, $F(1, 65) = 30.8$, $p < .001$, and reviewing homework, $F(1, 65) = 27.8$, $p < .0005$, whereas TAU therapists scored higher on other cognitive, $F(1, 65) = 16.1$, $p < .0005$, and general relaxation, $F(1, 65) = 49.0$, $p < .0005$, interventions. Finally, TAU therapists used a greater frequency of other interventions not specific to PCT or CBT in general, $F(1, 65) = 43.1$, $p < .0005$.

Discussion

We designed this study as a test of the effectiveness of PCT (Craske et al., 1994) compared with TAU in a capitated managed care system. Clients were relatively heterogeneous with a range of comorbid diagnoses and medication use prior to treatment. Therapists were primarily master's-level clinicians with little or no prior exposure to PCT or other cognitive–behavioral approaches. Overall, the results provide tentative support for the ability of PCT to enhance client outcomes compared with usual care. Within the intent-to-treat sample, clients treated by PCT-trained therapists showed significantly greater pre- to posttreatment change on measures of panic severity, depression, and general well-being than did clients receiving TAU.

Table 4
Mean Total Scores for PCT, General CBT, and Other Interventions in Both Treatment Conditions

Intervention	PCT		TAU		
	M	*SD*	*M*	*SD*	*F(df)*
PCT					
Education about panic	4.5	2.5	1.3	1.4	
Causes/maintenance of panic	4.6	2.5	0.7	1.1	
Specific goals of PCT	3.9	2.3	0.4	0.8	
Decatastrophizing	3.2	2.7	0.9	1.2	
Probability overestimation	3.3	0.3	0.2	0.8	
Prediction testing	0.8	1.3	0.0	0.2	
Interoceptive exposure	2.1	2.6	0.0	0.0	
Agoraphobic exposure	0.9	1.3	1.1	2.0	
Breathing retraining	3.1	2.5	1.3	1.5	
Self-monitoring	4.2	2.6	0.8	1.5	
Sequencing	2.4	2.3	0.1	0.6	
M	3.1	1.5	0.6	0.5	$F(1, 65) = 84.1$***
General CBT interventions					
Assigning homework	9.7	4.2	4.6	3.3	
Reviewing homework	5.6	4.0	1.5	2.2	
General relaxation	0.3	0.9	2.1	2.3	
Other cognitive techniques	1.9	2.3	5.4	4.4	
M	4.4	2.2	3.4	1.9	$F(1, 65) = 3.9$*
Other interventions					
Past experience	0.0	0.0	1.7	1.7	
Other treatment referrals	0.0	0.2	1.1	1.7	
Other specific treatment goals	0.2	0.7	2.3	2.3	
Assessing therapeutic progress	7.4	4.0	9.2	4.2	
Focus on therapeutic relationship	0.0	0.0	0.4	1.0	
Focus on interpersonal issues	0.9	1.7	5.2	4.1	
Assess reaction to treatment rationale	1.0	1.4	0.5	1.0	
M	1.3	0.7	2.9	1.2	$F(1, 65) = 43.1$***

Note. PCT = panic control therapy; CBT = cognitive–behavioral therapy; TAU = treatment as usual.
* $p < .05$. *** $p < .001$.

Our efforts to increase the generalizability of previous efficacy research might be expected to have weakened treatment effects compared with settings that include highly trained and supervised therapists, ample time to complete the entire treatment, and relatively homogeneous client samples. However, clients treated with PCT in the current study demonstrated decreases in panic severity comparable with clients treated in controlled efficacy trials (e.g., Barlow et al., 2000). Intent-to-treat cases in Barlow et al.'s study had an average posttest PDSS mean item score of .76, which was identical to mean posttreatment scores in the current study. These findings are consistent with recent benchmarking studies supporting the generalizability of findings from PCT efficacy studies to clinical practice settings (Wade et al., 1998; Warren, 1995). Thus, there is mounting evidence that concerns about the inability of ESTs to achieve positive outcomes in clinical practice may be unfounded (Fensterheim & Raw, 1996; Garfield, 1996; Silverman, 1996).

PCT produced greater change than TAU on several measures within both the intent-to-treat sample and the sample of treatment completers. There were no differences between the treatments among treatment noncompleters. This pattern of findings likely resulted from treatment completers carrying a significant portion of the between-treatments effect in the intent-to-treat sample. In addition, greater statistical power in the intent-to-treat versus the treatment noncompleter sample could have produced significant differences among the former but not the latter. A conservative interpretation of these findings is that, following treatment, mental health consumers receiving PCT in clinical practice are likely to be functioning better than if they had received usual care—provided that they remain in treatment for at least eight sessions. This is a critical consideration given the increase in session limits and other treatment-related constraints in managed care settings; in the current study, 60% of clients completed fewer than eight sessions in 8.5 months of treatment. Finally, it is important to note that the pattern of significantly greater change in PCT held for both panic-specific and more general measures of well-being including depression and overall functioning. These findings are consistent with other studies showing collateral improvement in comorbid conditions with empirically supported treatment for specific disorders (e.g., Brown, Antony, & Barlow, 1995).

Clinical significance analyses indicated varying rates of reliable and meaningful change from pre- to posttreatment for clients in both PCT and TAU. The percentage of clients achieving clinically significant change varied with measures of different anxiety-related problems. The highest rates of change were found for the severity of panic symptoms. These ranged from 70% of treatment completers in PCT to 42% of completers in TAU. Although there were consistent trends suggesting that PCT produced higher rates of clinically significant change than TAU, none of these analyses reached conventional levels of statistical significance.

Therapist adherence analyses indicated that the treatments were well discriminated. PCT therapists used more PCT interventions than TAU therapists, and TAU therapists were more likely to use other interventions such as focusing on interpersonal problems or discussing family-of-origin issues. However, as we expected, the discrimination was not perfect because PCT therapists periodically used interventions not specific to the treatment, and TAU therapists occasionally used PCT techniques. The results suggest that when PCT therapists used other techniques, the techniques tended to be general cognitive–behavioral interventions rather than more psychodynamic or interpersonally oriented strategies. In contrast, TAU therapists frequently used the latter strategies as well as some general cognitive–behavioral interventions not specific to PCT (e.g., challenging depressive thinking).

In general, the adherence data do not suggest a significant degree of overlap between the treatments. It seems unlikely that therapists in clinical practice will obtain levels of adherence comparable with those found in tightly controlled efficacy studies. Thus, the adherence results are somewhat predictable given the study's focus on the effectiveness of an empirically supported treatment as it would be delivered by clinical practitioners. Nonetheless, it should be noted that we found a low use of agoraphobic exposure techniques and no differences between the treatments. On the basis of conversations with therapists, our impression is that the practice setting (e.g., session limits, large caseloads) made it difficult to find the time necessary to accompany clients on predetermined exposure exercises. Given the high rate of agoraphobic avoidance in the current client sample, this may have reduced the likelihood of finding differences between the treatments. Considering the important role that in vivo agoraphobic exposure techniques play in obtaining positive treatment outcomes (Gould et al., 1995; White & Barlow, 2002), therapists' ability to use them in clinical practice may be a significant obstacle to disseminating PCT. Finally, we did not assess therapist competence in delivering PCT, and thus, we cannot determine how skillfully PCT therapists used the treatment; we can note only that they used a wide range of PCT techniques much more frequently than did TAU therapists.

The significant changes that occurred in the TAU condition are encouraging and support the continued study of usual care in clinical practice settings. If these results generalize to other practice settings, usual care will be an important focus for research to determine the range of effective interventions (empirically supported or otherwise) currently being used in clinical practice. Assessment of longer term client outcomes is one obvious next step and is currently underway in our research program. There are also other important outcomes to consider in addition to clinical status, including use of collateral services (e.g., emergency room, primary care, and specialty care visits), cost of treatments, and consumer and provider satisfaction. Each of these factors has been hypothesized to influence the effectiveness of ESTs in clinical practice and none yet have been tested with adequate controls (Addis et al., 1999; Addis & Waltz, 2002; Hatgis et al., 2001).

There are some limitations to the current study that need to be considered when one is interpreting the findings. Given the absence of a no-treatment or wait-list control condition, we cannot unequivocally attribute change to active treatment ingredients per se. However, findings from previous controlled efficacy studies suggest that PCT's effects are not solely due to nonspecific factors or regression to the mean (Gould et al., 1995; White & Barlow, 2002). Given the difficulty of implementing adequate control groups in clinical practice settings, it is likely that the results of efficacy studies will continue to play a role in the interpretation of treatment effectiveness outcomes.

It is also possible that nonspecific effects of therapist training were responsible for differences between PCT and TAU. Therapists conducting PCT received training and optional supervision. Therapists may have gained important clinical knowledge not specific to PCT as part of their training. Future research might consider including a therapist-training control condition. Relat-

edly, it is possible that expectancy effects played a role given that the principal investigator of the study ran the optional PCT supervision meetings and expressed confidence in the treatment. It is important to note that statistically and clinically significant changes were found across both anxiety-related problems and general well-being on both self-report measures and structured interviews. If expectancy effects were the primary mechanisms of change, the results would have been more likely to be restricted to panic-specific problems measured by self-report.

At the start of treatment, approximately one half of clients were currently using antianxiety medications, and one third were using antidepressant medication. Although pretreatment use was not correlated with treatment outcome in either condition, it is possible that beginning or continuing use during treatment carried treatment effects in both conditions that worked against finding differences. Unfortunately, we were unable to collect reliable data on medication use during treatment, although future analyses of pharmacy records may partially address medication effects. Nonetheless, medication treatment for anxiety disorders is common in clinical practice and provides part of the context into which ESTs will be disseminated. Thus, assessing psychosocial treatment effects for clients taking psychotropic drugs is an important aspect of effectiveness research (Addis & Waltz, 2002).

Although this study represents to our knowledge one of the first attempts to assess the effects of disseminating an EST to a managed care practice setting, the generalizability of the results to other practice contexts remains an open question. Our sample of 13 therapists was relatively small for the purposes of generalizing to large populations of practitioners. In addition, roughly half of the practitioners chose not to participate in the study. Incentive systems in managed care environments may be necessary to encourage practitioners to use empirically supported treatments. In the current study, treatment took place within a specialty clinic that had carved out mental health services for a large capitated HMO. Aspects of this practice context (e.g., session limits, large therapist caseloads, productivity pressures) may have affected the therapists' abilities to implement PCT. These are precisely the sort of clinical practice realities that must be considered in determining the best strategies to disseminate ESTs (Addis, 2002; Hatgis et al., 2001).

References

Addis, M. E. (1997). Evaluating the treatment manual as a means of disseminating empirically validated psychotherapies. *Clinical Psychology: Science and Practice, 4,* 1–11.

Addis, M. E. (2002). Methods for disseminating research products and increasing evidence-based practice: Promises, obstacles, and future directions. *Clinical Psychology: Science and Practice, 9,* 381–392.

Addis, M. E., & Krasnow, A. D. (2000). A national survey of practicing psychologists' attitudes toward psychotherapy treatment manuals. *Journal of Consulting and Clinical Psychology, 68,* 331–339.

Addis, M. E., Wade, W., & Hatgis, C. (1999). Barriers to evidence based practice: Addressing practitioners' concerns about manual based psychotherapies. *Clinical Psychology: Science and Practice, 6,* 430–441.

Addis, M. E., & Waltz, J. (2002). Implicit and untested assumptions about the role of psychotherapy treatment manuals in evidence based mental health practice. *Clinical Psychology: Science and Practice, 9,* 435–438.

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

Barlow, D. H., & Craske, M. G. (1989). *Mastery of your anxiety and panic.* Albany, NY: Graywind Publications.

Barlow, D. H., Gorman, J. M., Shear, M. K., & Woods, S. W. (2000). Cognitive–behavioral therapy, imipramine, or their combination for panic disorder: A randomized controlled trial. *Journal of the American Medical Association, 283,* 2529–2536.

Beck, A. T., Steer, R. A., & Garbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review, 8,* 77–100.

Beck, A. T., Ward, C., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 4,* 561–571.

Borkovec, T. D., & Nau, S. D. (1972). Credibility of analogue therapy rationales. *Journal of Behavior Therapy and Experimental Psychiatry, 3,* 257–260.

Brown, T. A., Antony, M. M., & Barlow, D. H. (1995). Diagnostic comorbidity in panic disorder: Effect on treatment outcome and course of comorbid diagnoses following treatment. *Journal of Consulting and Clinical Psychology, 63,* 408–418.

Brown, T. A., DiNardo, P. A., & Barlow, D. H. (1994). *Anxiety Disorders Interview Schedule for DSM–IV: Client interview schedule* (Adult version). Albany, NY: Graywind Publications.

Chambless, D. L., & Hollon, S. D. (1998). Defining empirically supported therapies. *Journal of Consulting and Clinical Psychology, 66,* 7–15.

Chambless, D. L., & Ollendick, T. H. (2000). Empirically supported psychological interventions: Controversies and evidence. *Annual Review of Psychology, 52,* 685–716.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (Rev. ed.). New York: Academic Press.

Cox, B. J., Parker, J. D. A., & Swinson, R. P. (1996). Confirmatory factor analysis of the Fear Questionnaire with social phobia patients. *British Journal of Psychiatry, 168,* 497–499.

Cox, B. J., Swinson, R. P., Parker, J. D., Kuch, K., & Reichman, J. T. (1993). Confirmatory factor analysis of the Fear Questionnaire in panic disorder with agoraphobia. *Psychological Assessment, 5,* 235–237.

Craske, M. G., Meadows, E., & Barlow, D. H. (1994). *Therapist's guide for the mastery of your anxiety and panic II & agoraphobia supplement.* Albany, NY: Graywind Publications.

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.

DiNardo, P. A., Moras, K., Barlow, D. H., Rapee, R. M., & Brown, T. A. (1993). Reliability of *DSM–III–R* anxiety disorder categories: Using the Anxiety Disorders Interview Schedule—Revised (ADIS–R). *Archives of General Psychiatry, 50,* 251–256.

Fensterheim, H., & Raw, S. D. (1996). Psychotherapy research is not psychotherapy practice. *Clinical Psychology: Science and Practice, 3,* 168–171.

Franklin, M. E., Abramowitz, J. S., Kozak, M. J., Levitt, J. T., & Foa, E. B. (2000). Effectiveness of exposure and ritual prevention for obsessive–compulsive disorder: Randomized compared with nonrandomized samples. *Journal of Consulting and Clinical Psychology, 68,* 594–602.

Garfield, S. L. (1996). Some problems associated with "validated" forms of psychotherapy. *Clinical Psychology: Science and Practice, 3,* 218–229.

Gillis, M. M., Haaga, D. A. F., & Ford, G. T. (1995). Normative values for the Beck Anxiety Inventory, Fear Questionnaire, Penn State Worry Questionnaire, and Social Phobia and Anxiety Inventory. *Psychological Assessment, 7,* 450–455.

Gould, R. A., Otto, M. W., & Pollack, M. H. (1995). A meta-analysis of treatment outcome for panic disorder. *Clinical Psychology Review, 15,* 819–844.

Hahlweg, K., Fiegenbaum, W., Frank, M., Schroeder, B., & von Witzleben, I. (2001). Short- and long-term effectiveness of an empirically supported treatment for agoraphobia. *Journal of Consulting and Clinical Psychology, 69,* 375–382.

Hamilton, M. (1967). Development of a rating scale for primary depressive illness. *British Journal of Social and Clinical Psychology, 6,* 276–296.

Hatgis, C., Addis, M. E., Zaslavsky, I., Jacob, K., Krasnow, A. D., DuBois, D., et al. (2001). Cross-fertilization vs. transmission: Recommendations for developing a bi-directional approach to psychotherapy dissemination research. *Applied and Preventive Psychology: Current Scientific Perspectives, 10,* 37–49.

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59,* 12–19.

Lambert, M. J., Hansen, N. B., Umphress, V., Lunnen, K., Okiishi, J., & Burlingame, G. M. (1996). *Administration and scoring manual for the OQ–45.2.* Stevenson, MD: American Professional Credentialing Services.

Marks, I. M., & Matthews, A. M. (1978). Brief standard self-rating for phobic patients. *Behavior Research and Therapy, 17,* 263–267.

Miller, G. A., & Chapman, J. P. (2001). Misunderstanding analysis of covariance. *Journal of Abnormal Psychology, 110,* 40–48.

Morgenstern, J., Blanchard, K. A., Morgan, T. J., Labouvie, E., & Hayaki, J. (2001). Testing the effectiveness of cognitive–behavioral treatment for substance abuse in a community setting: Within treatment and posttreatment findings. *Journal of Consulting and Clinical Psychiatry, 69,* 1007–1017.

Mueller, R., Lambert, M. J., & Burlingame, G. (1998). The Outcome Questionnaire: A confirmatory factor analysis. *Journal of Personality Assessment, 70,* 248–262.

Overall, J. E., & Woodward, J. A. (1977). Nonrandom assignment and the analysis of covariance. *Psychological Bulletin, 84,* 588–594.

Persons, J. B., Bostrom, A., & Bertagnolli, A. (1999). Results of randomized controlled trials of cognitive therapy for depression generalize to private practice. *Cognitive Therapy & Research, 23,* 535–548.

Persons, J. B., & Silberschatz, G. (1998). Are the results of randomized controlled trials useful to psychotherapists? *Journal of Consulting and Clinical Psychology, 66,* 126–135.

Schoenwald, S. K., & Hoagwood, K. (2001). Effectiveness, transportability, and dissemination of interventions: What matters when? *Psychiatric Services, 52,* 1190–1197.

Seggar, L. B., Lambert, M. J., & Hansen, N. B. (2002). Assessing clinical significance: Application to the Beck Depression Inventory. *Behavior Therapy, 33,* 253–269.

Shear, M. K., & Maser, J. D. (1994). Standardized assessment for panic disorder research: A conference report. *Archives of General Psychiatry, 51,* 346–354.

Shear, M. K., Ricci, P., Williams, J., Frank, E., Grochocinski, V., Vander Bilt, J., et al. (2001). Reliability and validity of the Panic Disorder Severity Scale: Replication and extension. *Journal of Psychiatric Research, 35,* 293–296.

Silverman, W. H. (1996). Cookbooks, manuals, and paint-by-numbers psychotherapy in the 90s. *Psychotherapy, 33,* 207–215.

Strosahl, K. D. (1998). The dissemination of manual-based psychotherapies in managed care: Promises, problems, and prospects. *Clinical Psychology: Science and Practice, 5,* 382–386.

Stuart, G. L., Treat, T. A., & Wade, W. A. (2000). Effectiveness of an empirically based treatment for panic disorder delivered in a service clinic setting: 1-year follow-up. *Journal of Consulting and Clinical Psychology, 68,* 506–512.

Turnkington, D., Kingdon, D., & Turner, T. (2002). Effectiveness of a brief cognitive–behavioural therapy intervention in the treatment of schizophrenia. *British Journal of Psychiatry, 180,* 523–527.

Tuschen-Caffier, B., Pook, M., & Frank, M. (2001). Evaluation of manual-based cognitive–behavioral therapy for bulimia nervosa in a service setting. *Behavioral Research and Therapy, 39,* 299–308.

Umphress, V. J., Lambert, M. J., Smart, D. W., Barlow, S. H., & Clouse, G. (1997). Concurrent and construct validity of the Outcome Questionnaire. *Journal of Psychoeducational Assessment, 15,* 40–55.

Wade, W. A., Treat, T. A., & Stuart, G. L. (1998). Transporting an empirically supported treatment for panic disorder to a service clinic setting: A benchmarking strategy. *Journal of Consulting and Clinical Psychology, 66,* 231–239.

Warren, R. (1995). Panic control treatment of panic disorder with agoraphobia and comorbid major depression: A private practice case. *Journal of Cognitive Psychotherapy, 9,* 123–134.

Warren, R., & Thomas, J. C. (2001). Cognitive–behavior therapy of obsessive–compulsive disorder in private practice: An effectiveness study. *Journal of Anxiety Disorders, 15,* 277–285.

Westen, D., & Morrison, K. (2001). A multidimensional meta-analysis of treatments for depression, panic, and generalized anxiety disorder: An empirical examination of the status of empirically supported therapies. *Journal of Consulting and Clinical Psychology, 69,* 875–899.

White, K. S., & Barlow, D. H. (2002). Panic disorder and agoraphobia. In D. H. Barlow (Ed.), *Anxiety and its disorders: The nature and treatment of anxiety and panic* (2nd ed., pp. 328–379). New York: Guilford Press.

Wilson, G. T. (1996). Manual based treatments: The clinical application of research findings. *Behavior Research and Therapy, 34,* 295–314.

Received September 9, 2002
Revision received August 12, 2003
Accepted September 25, 2003 ■