

## Transporting an Empirically Supported Treatment for Panic Disorder to a Service Clinic Setting: A Benchmarking Strategy

Wendy A. Wade  
Center for Behavioral Health

Teresa A. Treat and Gregory L. Stuart  
Indiana University

This work examines the transportability of cognitive-behavioral therapy (CBT) for panic disorder to a community mental health center (CMHC) setting by comparing CMHC treatment outcome data with the results obtained in two controlled efficacy trials. Participants were 110 clients with a primary diagnosis of panic disorder with or without agoraphobia; clients were not excluded on the basis of medication use or changes, severity or frequency of panic attacks, age, or the presence of agoraphobia. Clients completed a 15-session CBT protocol. Despite differences in settings, clients, and treatment providers, the treatment outcomes for clients completing treatment in the CMHC and the efficacy studies were comparable: Of the CMHC clients who completed treatment, 87% were panic-free at the end of treatment, and clients showed significant reductions in anticipatory anxiety, agoraphobic avoidance, generalized anxiety, and symptoms of depression. The present study suggests that panic control treatment can be transported to a CMHC. Challenges facing the transportability of research-based treatment to CMHC clients, settings, and treatment providers are discussed.

Controlled treatment outcome research commonly takes place under conditions that maximize both internal validity and the specificity of conclusions about causal mechanisms. It can be argued that outcome findings from internally valid efficacy studies cannot be assumed to generalize to other settings, populations, and treatment providers (Hollon, 1996; Jacobson & Christensen, 1996; Seligman, 1996). How well do the results of empirically supported treatments (ESTs; see Jacobson & Christensen, 1996) hold up in natural settings? Generalizability studies are needed to determine the *transportability* of ESTs to community settings (Wilson, 1995, 1996).

Some consider efficacy research overvalued because of limitations in generalizability and have proposed the “effectiveness study” as a preferable alternative (Hoagwood, Hibbs, & Brent, 1995). The effectiveness study, in which efficacious interventions are examined in real-world service settings, is offered as a viable solution to the shortcomings of the traditional, controlled conditions of efficacy research. Proponents argue that effectiveness study methodology can be used to determine treatment effectiveness without sacrificing generalizability to actual clinical settings.

Unfortunately, the effectiveness study method is not without its problems. As critics have pointed out, the effectiveness ap-

proach compromises internal validity, thereby limiting conclusions of treatment effectiveness, particularly the causal factors of therapeutic change (Hollon, 1996; Jacobson & Christensen, 1996). As long as the question “What treatments work and for whom?” (Kiesler, 1966) is considered important, there is a critical need for efficacy studies. The question of what treatments work is best addressed by using the controlled trials of efficacy studies. The question of for whom such treatments work may require an additional step: the generalizability study. The combined results from efficacy and effectiveness studies can provide answers to both of these critical questions.

We adopted the “benchmarking” research strategy (McFall, 1996) to assess the transportability of cognitive-behavioral therapy (CBT) of panic disorder (Barlow, Craske, Cerny, & Klosko, 1989; Klosko, Barlow, Tassinari, & Cerny, 1994; Margraf, Barlow, Clark, & Telch, 1993; Telch et al., 1993) to a community mental health center (CMHC) setting. The benchmarking strategy uses point-by-point comparisons of the treatment outcome data obtained in research clinic settings with outcome data obtained in clinical service settings. This strategy allows us to determine whether the magnitude of change across various dimensions of a disorder is similar in research settings and service clinics. In essence, we use the magnitude of change obtained in efficacy studies as a benchmark against which to judge the magnitude of change in service clinic settings.

The benchmarking strategy is a variant of the replication study in which the generalizability of ESTs can be evaluated. Benchmarking yields important feasibility and effectiveness data from clinical service settings. In turn, the results of a benchmarking study provide critical feedback to efficacy researchers who, on the basis of these generalizability data, can better understand the transportability of ESTs to uncontrolled treatment settings.

Panic disorder, which occurs in 1.5% to 3.5% of the population (Markowitz, Weissman, Ouellette, Lish, & Klerman, 1989),

---

Wendy A. Wade, Center for Behavioral Health, Bloomington, Indiana; Teresa A. Treat, Department of Psychology, Indiana University; Gregory L. Stuart, Department of Psychology, Indiana University.

This study would not have been possible without the support of the board of directors, administration, and staff of the Center for Behavioral Health (CBH). We thank all the CBH staff who worked to make this program a success. We also thank Richard M. McFall, Richard J. Viken, and David H. Barlow for their invaluable expertise and advice.

Correspondence concerning this article should be addressed to Wendy A. Wade, CBH, 645 South Rogers Street, Bloomington, Indiana 47403-2367.

is associated with extensive negative social and health consequences. Specifically, researchers have demonstrated that people with panic disorder were at increased risk for substance abuse; suicide attempts; impaired social, marital, and vocational functioning; and physical and emotional health problems, resulting in greater use of medical and psychiatric care and increased use of psychoactive medications (Markowitz et al., 1989).

Data from research clinics provide a strong case for the use of CBT in the treatment of panic disorder. Several reviews of the empirical literature, including meta-analyses, have shown that CBT produces superior treatment outcomes relative to pharmacotherapy (Clum, 1989; Clum, Clum, & Surls, 1993; Gould, Otto, & Pollack, 1995; Michelson & Marchione, 1991). Moreover, CBT is associated with lower relapse rates and lower rates of attrition than drug treatments (Clum, 1989; Gould et al., 1995; Michelson & Marchione, 1991). Over the long term, group-administered CBT also is less expensive than pharmacotherapy (Gould et al., 1995). On average, 80–90% of clients treated with CBT are panic-free at the end of treatment (Barlow, Craske, Cerny, & Klosko, 1989; Klosko et al., 1994; Margraf et al., 1993; Telch et al., 1993; for meta-analytic reviews, see Chambless & Gillis, 1993; Clum et al., 1993; Gould et al., 1995). In addition, the treatment effects are enduring. At follow-up intervals ranging from 6 months to 2 years, 75–87% of clients who received CBT were panic-free (e.g., Brown & Barlow, 1995; Craske, Brown, & Barlow, 1991).

Normative comparisons, in which treated individuals are compared with nondisturbed individuals, are helpful in evaluating the clinical significance or value of therapeutic interventions (Kendall & Grove, 1988). Barlow et al. (1989), Brown and Barlow (1995), Craske et al. (1991), and Telch et al. (1993) used composite indices of high endstate status to measure the proportion of participants who fell within the normative range of functioning posttreatment. Although the individual measures that made up these respective composite indices differed across studies, the composites included a combination of client self-reported panic attacks and other measures. The results of high endstate status analyses have varied from study to study: Overall, 36% to 83% of clients receiving CBT for panic disorder maintain high endstate status between posttreatment and 24-month follow-up (Barlow et al., 1989; Brown & Barlow, 1995; Craske et al., 1991; Klosko et al., 1994; Telch et al., 1993).

Assessment of the transportability of CBT for panic disorder is overdue: Panic disorder seriously interferes with the lives of millions of people and a promising EST has been identified. To address this need, we used two efficacy studies (Barlow et al., 1989; Telch et al., 1993) as benchmarks against which to compare the CMHC outcome results. Exclusionary criteria were minimized, paralleling the efforts of Persons and colleagues in the study of depression (Persons, Burns, & Perloff, 1988; Persons, 1995). Unlike many efficacy studies, no adult with a primary diagnosis of panic disorder was excluded from treatment on the basis of age, comorbid diagnosis, medical problems, treatment history, use of medications, or personality dysfunction (Barlow et al., 1989; Gould et al., 1995; Otto & Pollack, 1994). In the current study, we chose an empirically supported, manualized CBT protocol and outcome measures that were feasible to gather in a CMHC setting and that compared adequately to

those of the benchmark efficacy studies. A brief description of the two relevant efficacy studies follows.

Barlow et al. (1989) compared participants in the following three treatment conditions with a wait-list control group: (a) applied progressive muscle relaxation; (b) interoceptive exposure, breathing retraining, and cognitive restructuring; and (c) muscle relaxation combined with interoceptive exposure, breathing retraining, and cognitive restructuring. The second treatment condition ultimately became the basis for panic control treatment (cf. Barlow & Craske, 1989, 1994). This second condition is the manualized treatment protocol administered in the present study. Barlow et al. (1989) excluded clients who were younger than age 18 and older than age 65, who initiated use of anxiolytic medications in the 3 to 6 months before treatment, and who exhibited moderate or severe agoraphobia. Barlow et al. (1989) treated clients in individual therapy format.

Telch et al. (1993) randomly assigned clients either to panic control treatment (Barlow & Craske, 1989) or to a delayed-treatment control group. Thus, for both relevant efficacy studies and the current benchmarking study, the same manualized treatment protocol was used. The exclusionary criteria in Telch et al.'s (1993) study were similar to those in Barlow et al.'s (1989) study in that they excluded individuals who were younger than age 18 and older than age 65, and they excluded clients with recent changes in psychotropic medications. Telch et al. also excluded persons who did not report any panic attacks within 30 days before treatment but did not exclude those with moderate and severe agoraphobia. Participants in the Telch et al. (1993) study were treated in group therapy format.

## Method

### Participants

Participants were 110 clients of the Center for Behavioral Health in Bloomington, Indiana. Clients were seeking treatment for panic disorder and were self-referred in response to advertisements (primarily yellow pages, newspaper, and radio advertisements) or were referred by physicians, mental health professionals, or community agencies. In most cases, referring primary care physicians transferred their patients' full treatment for anxiety, including management of psychotropic medication, to the CMHC clinic.

Clients with a primary *Diagnostic and Statistical Manual of Mental Disorders*, third edition, revised (*DSM-III-R*; American Psychiatric Association, 1987) diagnosis of panic disorder with agoraphobia (PDA) or without agoraphobia (PD) were admitted for treatment. Exclusionary criteria included active symptoms of alcohol or drug dependency, psychosis, or mental disorder caused by a medical condition; these exclusionary criteria also were employed in the Barlow et al. (1989) and Telch et al. (1993) studies. In most cases, because of the time-intensive nature of the panic control treatment, clients concurrently involved in other psychotherapy were not admitted into treatment but were asked to return after completion of their ongoing treatment. Unlike some of the original controlled studies of panic disorder (e.g., Brown & Barlow, 1995; Barlow et al., 1989; Telch et al., 1993), no exclusions were made on the basis of medication use or changes, severity or frequency of panic attacks, age, or the presence or severity of agoraphobia.

### Measures

*Diagnostic interview.* *DSM-III-R* diagnoses were based on a modified version of the Anxiety Disorders Interview Schedule—Revised

(ADIS-R; DiNardo & Barlow, 1988). This version omitted research questions unnecessary to arrive at a diagnosis. The ADIS-R is a semi-structured interview with well-established psychometric properties (e.g., Barlow et al., 1989; Brown, Antony, & Barlow, 1995; Brown & Barlow, 1995; Craske et al., 1991; Klosko et al., 1994; Shear, Pilkonis, Cloitre, & Leon, 1994). The ADIS-R assesses the *DSM-III-R* anxiety disorders and mood disorders and screens for other major disorders (e.g., substance abuse, psychosis, or somatoform disorders). ADIS-R interviews were conducted by psychologists, master's-level clinicians, and advanced clinical psychology graduate students, all of whom received extensive training in the use and scoring of the ADIS-R. Each case was reviewed by the clinic director at weekly multidisciplinary team meetings of psychologists, a psychiatrist, master's-level clinicians, and graduate students. Diagnostic issues stemming from problems such as client inconsistencies in self-reporting and the presence of medical conditions that might contribute to symptomatology often necessitated gathering additional information. In such cases, diagnoses were determined jointly, after additional information and physician input were collected.

**Self-report data.** A battery of self-report questionnaires was administered before, at the middle of, and end of treatment. The end-of-treatment assessment was completed at the conclusion of the last therapy session. In the Results and Discussion sections, end of treatment is referred to as posttreatment.

The Fear Questionnaire (FQ; Marks & Mathews, 1979), a 15-item self-report questionnaire, was designed to assess phobic avoidance. The FQ contains three five-item subscales: Agoraphobia (FQ-Ag), Blood/Injury (FQ-BI), and Social Phobia (FQ-SP). Marks and Mathews reported adequate psychometric properties for each of the FQ subscales. Scores of less than 12 on each subscale are considered normative. The FQ is the most commonly used measure for assessing agoraphobia in treatment outcome research (Jacobson, Wilson, & Tupper, 1988).

The Beck Depression Inventory (BDI; Beck, Rush, Shaw, & Emery, 1979) is a 21-item self-report questionnaire used to assess depressive symptoms (Beck, Steer, & Garbin, 1988). Scores of 10 or less are considered normative. Beck et al. (1988) reviewed numerous studies conducted over the past 25 years documenting the concurrent, discriminant, and construct validity of the BDI.

The Positive and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988) contains two 10-item mood scales: positive affect (PA) and negative affect (NA). Clients rated the extent to which they experienced each emotion on the PANAS in the past week. Watson et al. (1988) reported adequate psychometric properties for the PANAS.

**Self-monitoring.** Once admitted into treatment, clients kept daily self-monitoring records of anxiety and other symptoms. All self-monitoring records are included in the panic control treatment client manual, *Mastery of Your Anxiety and Panic* (MAP; Barlow & Craske, 1994). Panic attack records, which include the date, time, duration, severity, symptoms, and setting in which the panic attack occurred (e.g., people who were present, expected or unexpected panic attack, etc.), were distributed to the clients. Clients were instructed to record the details of each panic attack as soon after the incident as possible, to maximize the accuracy of their reports. For normative comparisons, the criterion was set at no panic attacks.

Clients also kept daily ratings of depressed mood, general anxiety, and anticipatory anxiety (i.e., worry about panicking). A 9-point Likert scale was used for the daily mood ratings, ranging from 0 (*none*) to 8 (*very severe*). Scores of 2 and less were considered normative; this cutoff is consistent with the criterion used by Barlow et al. (1989) and Craske et al. (1991). Daily ratings were averaged for each week of treatment. In this study, we present self-monitoring data for the week before the first panic group and the last week of treatment.

### Procedure

Clients contacted the center between Fall 1992 and Fall 1995 and were prescreened for presence or absence of anxiety symptoms before

diagnostic interviews were conducted. The majority (82.7%) of clients were treated in group format, 11.8% of the clients were treated individually, and 5.5% of the clients received a combination of group and individual treatment. Individual treatment was provided to clients who had schedule conflicts with group meeting times, who displayed the potential for being disruptive in group, or who might have had difficulty keeping up with the pace of the group because of poor comprehension abilities, including limited reading skills.

Clients participated in panic control treatment, a 15-session CBT protocol described in the MAP treatment manual (Barlow & Craske, 1994). The treatment included psychoeducation regarding panic disorder, cognitive restructuring, diaphragm breathing retraining, interoceptive exposure, and naturalistic exposure. In addition, clients with a diagnosis of agoraphobia received a two-session module on agoraphobic exposure (Barlow & Craske, 1994). Session 1 of the protocol was conducted individually for each participant, and significant others (e.g., spouses, partners, parents) were encouraged to attend. Sessions 2 to 14 consisted of weekly, 90-min group sessions. The majority of clients treated in a group format received an optional individual therapy session after Session 9 to ensure that cognitive interventions were utilized appropriately. Clients who were treated individually attended weekly, 60-min sessions. The two treatment sessions that dealt with agoraphobic exposure were optional for those without an agoraphobia diagnosis; thus, clients who declined these optional sessions received 12 or 13 sessions. For those participants who met the criteria for panic disorder with agoraphobia, significant others were invited to attend the participants' first agoraphobia exposure session so that they could participate as coaches in this phase of treatment.

Participation in treatment required substantial client involvement and motivation. For example, participants were expected to complete daily self-monitoring, weekly readings in the client MAP manual, diaphragm breathing practices, thought identification and modification, and exposure practices.

**Treatment providers.** Primary therapists were psychologists and master's level clinicians self-selected from the CMHC adult outpatient staff. Training and supervision were extensive to maximize therapist adherence to the treatment protocol. The general strategy was a preceptor model. The clinic director received advanced training and certification in panic control treatment, and she, in turn, trained the staff. Training procedures included reading theoretical articles and treatment outcome research; learning the treatment manuals (therapist and client); participating in weekly seminarlike discussions about diagnosis, treatment, and theoretical issues; viewing videotapes and listening to audiotapes of actual panic control treatment sessions; and direct observation of the clinic director and practice with corrective feedback.

Therapist adherence to the treatment was emphasized throughout training and service delivery. For example, every group session was led by a senior therapist who was proficient in the treatment protocol. Therapists in training observed at least one group, from start to finish, before taking an active role in treatment. As they assumed an active role in treatment, therapists in training were directly observed by a senior therapist, who provided corrective feedback. Supervision of individual cases consisted of case reviews and feedback on audiotaped sessions. All individual sessions provided by junior therapists were audiotaped and closely supervised by the clinic director. In addition, to further encourage treatment adherence, structured outlines containing the information to be covered at each session were provided to therapists. Our training strategies resemble those discussed by Kendall and Southam-Gerow (1995) and Weisz, Donnemberg, Han, and Weiss (1995).

Psychopharmacological services were provided for clients who entered the program already medicated; who presented with physical or medical conditions that might cause, contribute to, or complicate psychological functioning; who reported severe depression symptoms that

Table 1  
*Descriptive Characteristics of Clients in CMHC, Barlow et al. (1989), and Telch et al. (1993) Samples Who Received Highly Similar Treatment*

Variable	CMHC	Barlow et al. (1989)	Telch et al. (1993)
Sample size	110	15	34
Mean age ( <i>SD</i> )	31.1 (9.1)	36.1 (6.8)	36.9 (11.7)
Gender (% female)	70.9	66.7	76.5
Marital status (%)			
Never married	36.7	—	32.3
Married	51.4	—	41.2
Divorced/separated	11.9	—	20.5
Education (%)			
Less than high school	10.0	—	0.0
High school completed	31.8	—	9.1
Partial college	29.1	—	42.4
College completed	29.1	—	48.5
No. of months ( <i>SD</i> ) since first panic	70.6 (82.3)	88.2 (69.8)	—
% using anxiolytic medication	56.9	40.0	47.1
% using depression medication	25.7	6.7	11.8
Comorbid Axis I diagnosis (%)	41.3	—	—
Agoraphobic (%)			
None	20.6	—	—
Mild	32.7	—	—
Moderate	30.8	0.0	—
Severe	15.9	0.0	—

Note. Dashes indicate that the information was not reported. CMHC = community mental health center.

might interfere with CBT; or whose initial anxiety was severe enough to interfere with CBT.

## Results

Results are presented in four sections. First, participant characteristics are described and compared with the Barlow et al. (1989) and Telch et al. (1993) samples. Second, the participants in our sample who completed treatment are compared with those who dropped out of treatment. Third, pre- to posttreatment changes in our sample are contrasted with the parallel benchmark results of Barlow et al. (1989) and Telch et al. (1993). Finally, predictors of treatment outcome are examined.

*Comparison of CMHC sample with samples from controlled efficacy studies.* Descriptive statistics from the CMHC, Barlow et al. (1989), and Telch et al. (1993) samples are presented in Table 1. The mean age of the CMHC sample is younger, and the age range is greater ( $M = 31.1$ ; range = 16–68 years) than those in the Barlow et al. sample ( $M = 36.1$ ; range = 18–65) and the Telch et al. sample ( $M = 36.9$ ; range = 18–65).<sup>1</sup>

Clients treated in the CMHC had fewer years of education, were more likely to be taking anxiolytic or antidepressant medication at pretreatment, and endorsed more severe distress on the agoraphobia subscale of the FQ as compared with the participants in the efficacy studies. Barlow et al. (1989) controlled for agoraphobia, whereas Telch et al. (1993) did not. Of the CMHC sample, 46.7% met criteria for moderate or severe agoraphobia.

A full listing of comorbid Axis I diagnoses from the CMHC sample is provided in Table 2.

*Comparison of completers with dropouts in the CMHC sample.* Of the 110 clients who initiated treatment, the 29 clients

(26.4%) who failed to attend at least 8 of the first 11 sessions were categorized as noncompleters. Of the 29 noncompleters, 4 discontinued panic control treatment and initiated depression treatment on the recommendation of the therapist. In all four of these cases, clients reported intensification of depressive symptoms and the presence of suicidal ideation. Two other patients moved before completing treatment. Four more dropped out because of work conflicts. Three people left treatment because of physical illnesses, and 2 discontinued for financial reasons. Of the 29 noncompleters, 14 left treatment for unknown reasons. Barlow et al. (1989) reported that only one participant dropped out of treatment; Telch et al. (1993) reported no attrition. Thus, statistical comparisons of completers and dropouts from the CMHC sample could not be made with completers and dropouts from the efficacy studies.

Attrition rates for the present sample were low in comparison with the 55% dropout rate reported in a study of interpersonal psychotherapy for depression conducted within the same CMHC setting between 1993 and 1995 (Lee, 1995). In a large-scale study of health service delivery, Phillips (1987) reported that 27% of clients receiving mental health services remained in treatment at 6 weeks and only 5.5% remained at 15 weeks. Thus, the attrition rate of the present sample is better than might be expected for this CMHC population.

<sup>1</sup> Brown et al. (1995) reported sample characteristics for a much larger sample ( $n = 126$ ). These data are very similar to those of the Barlow et al. (1989) benchmarking study and to those of the current study. The number of months since first panic (66.44) approaches the results of the current study. Also similar to the present study, comorbid Axis I diagnoses occurred in 51% of the Brown et al. (1995) sample.

Table 2  
Frequencies (and Percentages) of Comorbid DSM-III-R Axis I  
Diagnoses for the CMHC Sample

Diagnosis	Noncompleters <sup>a</sup>	Completers <sup>b</sup>	Total
Generalized anxiety disorder	10 (34.4%)	13 (16.3%)	23 (21.2%)
Major depression	8 (27.6%)	14 (17.5%)	22 (20.2%)
Posttraumatic stress disorder	1 (3.4%)	3 (3.8%)	4 (3.7%)
Social phobia	1 (3.4%)	2 (2.5%)	3 (2.8%)
Simple phobia	1 (3.4%)	1 (1.3%)	2 (1.8%)
Obsessive-compulsive disorder	0 (0.0%)	2 (2.5%)	2 (1.8%)
Dysthymia	2 (6.9%)	0 (0.0%)	2 (1.8%)
Adjustment disorder	1 (3.4%)	1 (1.3%)	2 (1.8%)
Partner relational problem	0 (0.0%)	2 (2.5%)	2 (1.8%)
Anorexia nervosa	1 (3.4%)	0 (0.0%)	1 (0.9%)
Bipolar disorder	0 (0.0%)	1 (1.3%)	1 (0.9%)

Note. DSM-III-R = Diagnostic and Statistical Manual of Mental Disorders (3rd ed., revised). CMHC = community mental health center.

<sup>a</sup>  $n = 29$ . <sup>b</sup>  $n = 80$ ; data were missing for 1 client.

Differences between completers and noncompleters were examined using independent  $t$  tests for continuous variables and chi-square tests for categorical variables. Comparisons between completers ( $n = 81$ ) and noncompleters ( $n = 29$ ) on pretreatment variables are summarized in Table 3. The sample size differs across variables because of missing data. There were no significant differences between completers and noncompleters on frequency of panic attacks, the three Fear Questionnaire subscales (i.e., FQ-Ag, FQ-BI, and FQ-SP), the PANAS Negative Affect subscale, and the proportion of clients taking anxiolytic medications.

There was a nonsignificant trend for more noncompleters to meet criteria for major depressive disorder (completers = 19.4%, noncompleters = 30.8%),  $\chi^2(1, N = 98) = 1.41, ns$ .

Group differences between completers and dropouts did emerge on some variables. Completers were significantly older and had received more years of education than noncompleters. In addition, completers reported longer standing panic and agoraphobic symptoms, fewer depressive symptoms, and less anti-depressant medication use at intake. Finally, completers were

Table 3  
Pretreatment Comparison of Noncompleters and Completers

Variable	Noncompleters	Completers	df	$t$ or $\chi^2$
Average no. panics in last week	1.73 (1.85)	1.99 (2.33)	99	0.51
Fear Questionnaire—Agoraphobia	19.28 (14.74)	16.04 (11.07)	99	-1.17
Fear Questionnaire—Blood/Injury	17.92 (9.66)	14.12 (8.11)	99	-1.94
Fear Questionnaire—Social Phobia	19.28 (12.06)	16.42 (8.90)	99	-1.27
Age at intake	27.66 (6.02)	32.28 (9.75)	108	2.39*
Education (%)			3	16.44***
Less than high school	27.6	3.7		
High school completed	37.9	29.6		
Partial college	17.2	33.3		
College completed	17.2	33.3		
Years since first panic attack	3.38 (3.62)	6.76 (7.50)	106	2.29*
Years since first agoraphobic episode	2.37 (2.99)	6.15 (6.14)	78	2.65**
PANAS Positive Affect score	22.9 (6.7)	27.5 (8.1)	103	2.61*
PANAS Negative Affect score	31.9 (7.8)	29.6 (9.2)	103	-1.14
Beck Depression Inventory score	20.2 (11.0)	14.6 (8.7)	106	-2.40*
Depression medication (% using)	41.4	20.0	1	5.10*
Anxiolytic medication (% using)	48.3	60.0	1	1.19
Comorbid Axis I diagnosis (%)			3	8.23*
No comorbid diagnosis	37.9	66.3		
One comorbid diagnoses	37.9	18.8		
Two comorbid diagnoses	24.1	13.8		
Three comorbid diagnoses	0.0	1.0		

Note. Values enclosed in parentheses are standard deviations. PANAS = Positive and Negative Affect Scale.

\*  $p < .05$ . \*\*  $p < .01$ . \*\*\*  $p < .001$ .

Table 4  
*Characteristics of Clients in CMHC, Barlow et al. (1989), and Telch et al. (1993)*  
*Samples Who Received Similar Treatment*

Variable	CMHC <sup>a</sup>		Barlow et al.		Telch et al.	
	Pre	Post	Pre	Post	Pre	Post
% panic-free	18.2	87.2	15.4	84.6	29.4	85.3
Mean no. panics in past week	2.0 (2.4)	0.1 (0.3)	1.2 (1.0)	1.0 (2.6)	4.2 (9.5)	0.2 (0.5)
Mean anticipatory anxiety	3.4 (1.5)	1.6 (1.3)	—	—	—	—
Mean general anxiety	3.6 (1.3)	2.4 (1.3)	1.9 (1.0)	2.0 (2.2)	—	—
Mean depression	2.4 (1.7)	1.3 (1.5)	1.7 (0.8)	1.3 (0.9)	—	—
FQ-Ag	16.5 (10.8)	8.8 (6.9)	—	—	12.2 (11.4)	5.1 (6.8)
FQ-SP	17.3 (8.9)	10.1 (6.5)	—	—	—	—
FQ-BI	14.9 (8.2)	10.7 (7.1)	—	—	—	—
BDI score	15.5 (8.8)	6.0 (6.3)	13.5 (8.9)	11.3 (7.2)	16.9 (8.2)	7.7 (5.3)
PANAS PA	26.0 (8.0)	32.4 (7.8)	—	—	—	—
PANAS NA	30.7 (8.3)	17.3 (6.3)	—	—	—	—
% using anxiolytic medication	60.0	23.1	40.0	13.0	47.1	47.1
% using depression medication	20.0	16.9	7.0	0.0	11.8	11.8

*Note.* Values enclosed in parentheses are standard deviations. Dashes indicate that the information was not reported. CMHC = community mental health center; Pre = pretreatment; Post = posttreatment; FQ-Ag = Agoraphobia subscale of Fear Questionnaire; FQ-SP = Social Phobia subscale of Fear Questionnaire; FQ-BI = Blood/Injury subscale of Fear Questionnaire; BDI = Beck Depression Inventory; PANAS = Positive and Negative Affect Scale; PA = positive affect; NA = negative affect.

<sup>a</sup> Because of missing data, the sample sizes vary on each measure.

significantly less likely to have one or more comorbid *DSM-III-R* Axis I diagnoses. In addition, completers were significantly less likely to meet comorbid *DSM-III-R* Axis I diagnosis of generalized anxiety disorder (completers = 18.1%, non-completers = 38.5%),  $\chi^2(1, N = 98) = 4.43, p < .04$ .

*Comparisons of treatment changes in CMHC and controlled efficacy study samples.* We examined within-subject changes from pre- to posttreatment using paired *t* tests for continuous variables and chi-square tests for categorical variables. Benchmarking comparisons are made on eight variables, including (a) panic-free status, (b) frequency of panic attacks, (c) daily self-monitoring of general anxiety, (d) daily self-monitoring of depression, (e) FQ-Ag score, (f) BDI score, (g) percentage of clients taking anxiolytic medications, and (h) percentage of clients taking depression medications. As shown in Table 4, the CMHC treatment completers improved on virtually every measure, and the magnitude of these improvements was comparable to those of the improvements reported in the controlled efficacy studies of Barlow et al. (1989) and Telch et al. (1993). Specifically, from pre- to posttreatment: (a) the proportion of clients reporting panic-free status increased from 18.2% to 87.2%; (b) the weekly mean frequency of panic episodes decreased from 2 to 0.1,  $t(73) = 6.94, p < .001$ ; (c) clients' self-reported daily generalized anxiety decreased,  $t(60) = 7.39, p < .001$ ; (d) clients' self-reported daily depression decreased,  $t(56) = 5.11, p < .001$ ; (e) FQ-Ag subscale scores improved significantly,  $t(52) = 6.05, p < .001$ ; (f) BDI scores decreased significantly,  $t(54) = 7.69, p < .001$ ; (g) the proportion of clients receiving anxiolytic medication (benzodiazepines) decreased from 60.0% to 23.1%; and (h) the proportion of clients receiving antidepressant medication decreased from 20.0% to 16.9%.

Significant changes in the CMHC pre-post measures not

reported in the two efficacy studies included significant drops in daily anticipatory anxiety,  $t(59) = 9.51, p < .001$ . The FQ-SP and BI subscales decreased significantly,  $t(52) = 5.52, p < .001$  and  $t(53) = 4.29, p < .001$ , respectively. Scores on the PANAS PA and NA subscales improved significantly,  $t(53) = -6.22, p < .001$  and  $t(53) = 11.71, p < .001$ , respectively.

The clinical significance of treatment gains was determined through normative comparisons (Kendall & Grove, 1988) by examining the proportion of clients who attained scores in the normative range on relevant clinical dimensions of panic disorder. Four clinical dimensions were examined: (a) panic attacks, (b) anticipatory anxiety (worry about having a panic attack), (c) agoraphobic avoidance, and (d) depression. The proportion of CMHC treatment completers who fell within the normative range of functioning posttreatment as well as comparisons of corresponding data from benchmark studies are presented in Table 5. The magnitude of therapeutic change on all variables is similar across studies. Comparisons of agoraphobic avoidance (FQ-Ag) between participants in the present sample and the Telch et al. (1993) sample suggest that participants in the CMHC sample reported more severe avoidance both pre- and posttreatment; however, the magnitude of improvement (percentage of normative posttreatment minus percentage of normative pretreatment) was similar for the two samples (CMHC = 27%; Telch et al. [1993] = 24%). The present investigation did not include the entire array of measures used in the high endstate composites of previous studies; thus, composite benchmark comparisons could not be made.

*Predictors of treatment outcome in the CMHC sample.* We used multiple regression analyses to determine predictors of the frequency of posttreatment panic attacks, levels of self-reported anticipatory anxiety, and FQ-Ag subscale scores. The following

Table 5  
*Percentage of CMHC, Telch et al. (1993), and Barlow et al. (1989) Samples Scoring in the Normal Range of Functioning at Pre- and Posttreatment*

Variable	Criterion for recovery	CMHC <sup>a</sup>		Telch et al.		Barlow et al.	
		Pre	Post	Pre	Post	Pre	Post
Panic attacks	panics = 0	18.2	87.2	29.4	85.3	15.4	84.6
Anticipatory anxiety	score < 2	22.9	80.0	—	—	—	—
	ASI < 27	—	—	32.3	97.1	—	—
Avoidance	FQ-Ag < 12	40.5	67.6	61.8	85.3	—	—
Depression	BDI < 10	23.6	78.2	23.5	64.7	—	—

*Note.* Recovery criteria for panic attacks, anticipatory anxiety, FQ-Ag, and BDI are based on well-accepted norms reported in the literature. Dashes indicate that the information was not reported. CMHC = community mental health center; Pre = pretreatment; Post = posttreatment; ASI = Anxiety Sensitivity Index (Reiss, Peterson, Gursky, & McNally, 1986); FQ-Ag = Agoraphobia subscale of Fear Questionnaire; BDI = Beck Depression Inventory.

<sup>a</sup> Because of missing data, the sample sizes vary on each measure.

pretreatment variables were included in all three multiple regression analyses: age, gender, use of anxiolytic medication, use of antidepressant medication, frequency of panic attacks, presence of agoraphobia, presence of comorbid *DSM-III-R* diagnoses, and prior history of mental health treatment. Only use of anxiolytic medication significantly predicted the frequency of posttreatment panic attacks ( $\beta = 0.26$ ),  $t(64) = 2.18$ ,  $p < .05$ . Only pretreatment FQ-Ag subscale scores significantly predicted posttreatment FQ-Ag subscale scores, ( $\beta = 0.42$ ),  $t(43) = 2.77$ ,  $p < .01$ . None of the variables significantly predicted posttreatment levels of anticipatory anxiety.

### Discussion

We used a benchmarking research strategy (McFall, 1996) to evaluate the effectiveness of CBT for panic when transported to an uncontrolled service clinic. Despite differences in settings, the treatment outcomes for clients completing treatment in the present study and the efficacy studies were similar. Of the CMHC clients who completed treatment, 87% were panic-free at the end of treatment. Treatment was associated with significant reductions in anticipatory anxiety, agoraphobic avoidance, generalized anxiety, and symptoms of depression. Normative comparisons (Kendall & Grove, 1988) were used, and percentages of clients achieving normative functioning posttreatment on measures of panic, anticipatory anxiety, and depression were found to be similar in the present and the benchmark samples. However, the present sample was more severely agoraphobic pre- and posttreatment. At the same time, a large proportion of CMHC clients succeeded in meeting the treatment goal of discontinuing potentially addictive anxiolytic medications.

Several challenges emerging from transporting research protocols used in research clinics to the clients, treatment providers, and setting of a CMHC are worthy of consideration. The differences between CMHC treatment completers and noncompleters raise important issues. Education level was significantly lower in the noncompleter group as compared both with our treatment completers and with participants in the original efficacy studies. Although we made attempts to provide additional support to our less educated clients (e.g., an audiotaped version of the

client manual for barely literate clients; individualized rather than group sessions), these clients tended not to complete treatment. Additional research is needed to better understand how to successfully meet the treatment needs of less educated persons. Participants who did not complete treatment were significantly more likely to have one or more comorbid conditions. This finding contrasts with the efficacy study of Brown et al. (1995), which did not find pretreatment comorbidity predictive of non-completer status. If the present finding is typical of service clinics, then additional research is necessary to better serve this multiproblem group.

Another topic deserving additional research attention is agoraphobia. Nearly half of the individuals treated in our service clinic met criteria for moderate or severe agoraphobia. At posttreatment, there were sufficient continued agoraphobic fears that 33% of completers scored within the nonnormative range on the FQ-Ag. As we have begun to collect 1-year follow-up data on this sample, we have noted that some agoraphobic clients have continued to make progress on their own and have achieved normal functioning; others have not progressed well. More research is needed to determine how best to meet the treatment needs of persons with moderate and severe agoraphobia. These questions merit answers, especially as managed care organizations try to determine and define the parameters of necessary treatment.

The only significant predictor of posttreatment panic for our sample was use of anxiolytic medications (primarily benzodiazepines) at pretreatment. This finding is consistent with results of a recent investigation, in which ongoing benzodiazepine use was associated with earlier relapse (Otto, Pollack, & Sabatino, 1996). Continued research into the role of anxiolytic medications, psychosocial treatments, and their combination is indicated.

Wilson (1995) noted that therapists in controlled efficacy studies generally are selected because of their competence, expertise, and extensive training. Similarly, the transportability of ESTs may hinge in part on selecting and training competent treatment providers (Kendall & Southam-Gerow, 1995; Weisz et al., 1995). How much training and supervision are required

to effectively administer ESTs in services clinics is an empirical question. However, we must not assume that clinicians are able to implement these protocols effectively without receiving advanced training and supervision (Wilson, 1995), nor can we assume that clinicians will continue to adhere to standardized treatment protocols in the absence of continued monitoring and supervision. Manualized treatment protocols provide helpful structure and standardization of treatment, but they do not substitute for the theoretical foundation and skills acquired through advanced training and supervision. The recruitment of qualified, interested staff is necessary, but it can be difficult in service clinic settings. We have found that staff who lack prior training in CBT sometimes struggle in their attempts to learn and implement this treatment modality. In our experience, many clinicians are ambivalent about providing structured, short-term treatment in which adherence to a protocol is required. Well-intentioned therapists inexperienced in the administration of manualized treatment may abandon the protocol and revert to more familiar ideographically based treatment approaches (Persons, 1995; Wilson, 1995). Additional structure and support, in the form of treatment session outlines and weekly team meetings, were standard practice in this transportability endeavor. Despite the resources devoted to training and treatment integrity, some staff did not succeed in learning and adhering to panic control treatment. An unrelated staffing issue that may be unique to service clinic settings is availability of psychiatry services that are congruent with the goals of ESTs (in our case, CBT for panic disorder).

A final aspect of the transportability of controlled clinical trials is the feasibility of conducting ongoing treatment outcome research in a service clinic setting such as a CMHC. Use of the benchmarking strategy (McFall, 1996) was a natural and practical method for evaluating the transportability of panic control treatment. Our treatment intervention and outcome measures were chosen for their comparability to the existing efficacy research. Use of existing efficacy studies as road maps for effectiveness studies can increase our knowledge about what works and for whom.

Despite the strengths of benchmarking, the current study is not without its limitations. Some of the more serious limitations center around issues of internal validity. Given that random assignment and treatment control conditions were not feasible in our service clinic setting, our results must be interpreted with caution. The current investigation would have benefitted from the addition of comparison or contrast groups against which treatment effects could be gauged. Feasible comparison groups for future benchmarking studies could include wait-listed clients, dropouts, or clients who received alternative treatments.

In most service clinic settings, therapist variables cannot be manipulated. In our setting it was difficult to evaluate therapist variables because of the methods by which clients were assigned to therapists, the small number of therapists, and the use of cotherapists. However, benchmarking studies in larger treatment settings could examine these questions. Future investigations would be improved by systematic evaluation of diagnostic reliability. Such evaluations could be built into training and supervision. For example, supervisors could complete reliability checks while observing or listening to tapes of diagnostic interviews. Another limitation of the current investigation is the absence of

formal measures of treatment integrity. This omission can easily be overcome by including brief rating scales that parallel those used in efficacy studies. Treatment adherence data would improve the quality of effectiveness studies, would provide valuable quality control feedback, and would enhance supervision.

Client compliance with data collection was not a feasibility problem. On the contrary, most clients willingly participated in the process. The majority of problems with missing data were due to therapist or support staff error. Before treatment began, all clients were educated about the importance of tracking their symptoms and progress. Data collection of self-monitoring was described as essential to successful treatment and was built into each session; every session began with an individualized review of weekly self-monitoring.

When designed with careful forethought and consultation with efficacy researchers, methodologically sound effectiveness studies have much to offer. As managed health care proceeds to the forefront and the emphasis on accountability and outcome increases, collection of outcome data is becoming routine. These data will be more meaningful if they are collected in the context of a well-defined evaluation strategy. The benchmarking strategy is an attractive option that meets the market demand for outcome results while simultaneously and systematically advancing our knowledge of what treatments work and for whom.

## References

- American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., rev.). Washington, DC: Author.
- Barlow, D. H. (1994). Psychological intervention in the era of managed competition. *Clinical Psychology: Science and Practice*, 1, 109–122.
- Barlow, D. H., & Craske, M. G. (1989). *Mastery of your anxiety and panic*. Albany, NY: Graywind.
- Barlow, D. H., & Craske, M. G. (1994). *Mastery of your anxiety and panic II*. Albany, NY: Graywind.
- Barlow, D. H., Craske, M. G., Cerny, J. A., & Klosko, J. S. (1989). Behavioral treatment of panic disorder. *Behavior Therapy*, 20, 261–282.
- Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). *Cognitive therapy of depression*. New York: Guilford Press.
- Beck, A. T., Steer, R. A., & Garbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review*, 8, 77–100.
- 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., & Barlow, D. H. (1995). Long-term outcome in cognitive-behavioral treatment of panic disorder: Clinical predictors and alternative strategies for assessment. *Journal of Consulting and Clinical Psychology*, 63, 754–765.
- Chambless, D. L., & Gillis, M. M. (1993). Cognitive therapy of anxiety disorders. *Journal of Consulting and Clinical Psychology*, 61, 248–260.
- Clum, G. A. (1989). Psychological interventions vs. drugs in the treatment of panic. *Behavior Therapy*, 20, 429–457.
- Clum, G. A., Clum, G. A., & Surls, R. (1993). A meta-analysis of treatments for panic disorder. *Journal of Consulting and Clinical Psychology*, 61, 317–326.
- Craske, M. G., Brown, T. A., & Barlow, D. H. (1991). Behavioral treatment of panic disorder: A two-year follow-up. *Behavior Therapy*, 22, 289–304.



- DiNardo, P. A., & Barlow, D. H. (1988). *Anxiety Disorders Interview Schedule-Revised (ADIS-R)*. Albany: State University of New York, Phobia and Anxiety Disorder Clinic.
- 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.
- Hoagwood, K., Hibbs, E., & Brent, D. (1995). Introduction to the special section: Efficacy and effectiveness in studies of child and adolescent psychotherapy. *Journal of Consulting and Clinical Psychology, 63*, 683–687.
- Hollon, S. D. (1996). The efficacy and effectiveness of psychotherapy relative to medications. *American Psychologist, 51*, 1025–1030.
- Jacobson, N. S., & Christensen, A. (1996). Studying the effectiveness of psychotherapy: How well can clinical trials do the job? *American Psychologist, 51*, 1031–1039.
- Jacobson, N. S., Wilson, L., & Tupper, C. (1988). The clinical significance of treatment gains resulting from exposure-based interventions for agoraphobia: A reanalysis of outcome data. *Behavior Therapy, 19*, 539–554.
- Kendall, P. C., & Grove, W. M. (1988). Normative comparisons in therapy outcome. *Behavioral Assessment, 10*, 147–158.
- Kendall, P. C., & Southam-Gerow, M. A. (1995). Issues in the transportability of treatment: The case of anxiety disorders in youths. *Journal of Consulting and Clinical Psychology, 63*, 702–708.
- Kiesler, D. J. (1966). Some myths of psychotherapy research and the search for a paradigm. *Psychological Bulletin, 65*, 110–136.
- Klosko, J. S., Barlow, D. H., Tassinari, R., & Cerny, J. A. (1994). A comparison of alprazolam and behavior therapy in treatment of panic disorder. *Journal of Psychotherapy Practice and Research, 3*, 166–179.
- Lee, C. L. (1995, August). *Brief treatment of depression: Factors influencing recovery rate*. Paper presented at the annual meeting of the American Psychological Association, New York, NY.
- Margraf, J., Barlow, D. H., Clark, D. M., & Telch, M. J. (1993). Psychological treatment of panic: Work in progress on outcome, active ingredients, and follow-up. *Behavior Research and Therapy, 31*, 1–8.
- Markowitz, J., Weissman, M., Ouellete, R., Lish, J., & Klerman, G. (1989). Quality of life in panic disorder. *Archives of General Psychiatry, 46*, 984–992.
- Marks, I. M., & Mathews, A. M. (1979). Brief standard self-rating for phobic patients. *Behaviour Research and Therapy, 17*, 263–267.
- McFall, R. M. (1996, July 1). *Consumer satisfaction as a way of evaluating psychotherapy: Ecological validity and all that versus the good old randomized trial* (panel discussion). 6th Annual Convention of the American Association of Applied and Preventative Psychology, San Francisco.
- Michelson, L. K., & Marchione, K. (1991). Behavioral, cognitive, and pharmacological treatments of panic disorder with agoraphobia: Critique and synthesis. *Journal of Consulting and Clinical Psychology, 59*, 100–114.
- Otto, M. W., & Pollack, M. H. (1994). Treatment strategies for panic disorder: A debate. *Harvard Review of Psychiatry, 2*, 166–170.
- Otto, M. W., Pollack, M. H., & Sabatino, S. A. (1996). Maintenance of remission following cognitive-behavior therapy for panic disorder: Possible deleterious effects of concurrent medication treatment. *Behavior Therapy, 27*, 473–482.
- Persons, J. B. (1995). Why practicing psychologists are slow to adopt empirically validated treatments. In S. C. Hayes, V. M. Follette, R. M. Dawes, & K. E. Grady (Eds.), *Scientific standards of psychological practice: Issues and recommendations* (pp. 141–157). Reno, NV: Context Press.
- Persons, J. B., Burns, D. D., & Perloff, J. M. (1988). Predictors of dropout and outcome in cognitive therapy for depression in a private practice setting. *Cognitive Therapy and Research, 12*, 557–575.
- Phillips, E. L. (1987). The ubiquitous decay curve: Service delivery similarities in psychotherapy, medicine, and addiction. *Professional Psychology: Research and Practice, 18*(6), 650–652.
- Reiss, S., Peterson, R. A., Gursky, D. M., & McNally, R. J. (1986). Anxiety sensitivity, anxiety frequency and the prediction of fearfulness. *Behaviour Research and Therapy, 24*, 1–8.
- Seligman, M. E. P. (1995). The effectiveness of psychotherapy: The *Consumer Reports* study. *American Psychologist, 50*, 965–974.
- Seligman, M. E. P. (1996). Science as an ally of practice. *American Psychologist, 51*, 1072–1079.
- Shear, M. K., Palkonis, P. A., Cloitre, M., & Leon, A. C. (1994). Cognitive-behavioral treatment compared with nonprescriptive treatment of panic disorder. *Archives of General Psychiatry, 51*, 395–401.
- Telch, M. J., Lucas, J. A., Schmidt, N. B., Hanna, H. H., Jaimez, T. L., & Lucas, R. A. (1993). Group cognitive-behavioral treatment of panic disorder. *Behaviour Research and Therapy, 31*, 279–287.
- Telch, M. J., Schmidt, N. B., Jaimez, T. L., Jacquin, K. M., & Harrington, P. J. (1995). Impact of cognitive-behavioral treatments on quality of life in panic disorder patients. *Journal of Consulting and Clinical Psychology, 63*, 823–830.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54*(6), 1063–1070.
- Weisz, J. R., Donnerberg, G. R., Han, S. S., & Weiss, B. (1995). Bridging the gap between laboratory and clinic in child and adolescent psychotherapy. *Journal of Consulting and Clinical Psychology, 63*, 688–701.
- Wilson, G. T. (1995). Empirically-validated treatments as a basis for clinical practice: Problems and prospects. In S. C. Hayes, V. M. Follette, R. M. Dawes, & K. E. Grady (Eds.), *Scientific standards of psychological practice: Issues and recommendations* (pp. 163–196). Reno, NV: Context Press.
- Wilson, G. T. (1996). Manual-based treatments: The clinical application of research findings. *Behaviour Research and Therapy, 34*, 295–314.

Received December 10, 1996

Revision received June 23, 1997

Accepted June 26, 1997 ■