The Psychotherapy Dose-Response Effect and Its Implications for Treatment Delivery Services

Nathan B. Hansen, Yale University School of Medicine
Michael J. Lambert, Brigham Young University
Evan M. Forman, Drexel University

To date, few studies have been published on the dose-response relationship, but there is general consensus that between 13 and 18 sessions of therapy are required for 50% of patients to improve. Reviewing the clinical trials literature reveals that in carefully controlled and implemented treatments, between 57.6% and 67.2% of patients improve within an average of 12.7 sessions. Using naturalistic data, however, revealed that the average number of sessions received in a national database of over 6,000 patients was less than five. The rate of improvement in this sample was only about 20%. These results suggest that patients, on average, do not get adequate exposure to psychotherapy, nor do they recover from illness at rates observed in clinical trials research.

Key words: **psychotherapy outcome, dose-response, effectiveness research, efficacy research, clinical significance.** ***[Clin Psychol Sci Prac 9:329–343, 2002]***

A large body of psychotherapy research, accumulated over more than 40 years, has convincingly demonstrated the general effectiveness of psychotherapy (Lambert & Bergin, 1994). Additionally, a second generation of psychotherapy research has begun to identify particular treatments that work more or less effectively in specific contexts and populations (Chambless et al., 1996; DeRubeis & Crits-Christoph, 1998; Roth & Fonagy, 1996). Moreover, in some cases, specific parameters have been established regarding the duration of treatment required to achieve meaningful change in the majority of patients. This wealth of information has the potential to greatly inform standard clinical practice. However, it appears that psychotherapy research findings have had little impact on everyday clinical practice (Kopta, Lueger, Saunders, & Howard, 1999; Nathan, Stuart, & Dolan, 2000).

Further, despite the recognized and frequently addressed importance of psychotherapy research informing public policy in regard to mental health issues (DeLeon, 1988, Klerman, 1983; Newman & Howard, 1986; Strupp, 1986), it seems that there are serious difficulties in bridging this gap, not only between researchers and practitioners but also between researchers and policy makers (Speer, 1994; Wiggins, 1992). In this day of accountability and managed health care, the need for clinical research to reach and affect policy makers becomes especially important. For example, concerns have been raised by DeLeon, Vandenbos, and Bulatao (1991) about the current practices of managed health care organizations. These concerns included (1) services are limited as a result of a physician gate-keeping process; (2) unrealistic limits are enforced on the number of sessions provided and/or the amount of money available for service; (3) the quality and appropriateness of provided mental health service may be inadequate; (4) consumers of mental health services are not provided full information about the services available and, in fact, are usually given positive information while negative aspects of services are withheld; and (5) utilization review is often carried out by unlicensed individuals with little training in mental health issues. Clinical research can have a direct impact on these areas of concern, but as yet has not been widely cited, even by those who complain about the untoward effects of managed care policies and practices.

Cognizant of the difficulty of interpreting and applying clinical research in everyday practice, researchers have

Address correspondence to Nathan B. Hansen, Yale University School of Medicine, Department of Psychiatry, The Consultation Center, 389 Whitney Avenue, New Haven, CT 06511. E-mail: nathan.hansen@yale.edu.

been working to produce methods for conducting and reporting research that have greater clinical utility (Nathan et al., 2000). Clinical significance methodology (Jacobson, Follette, & Revenstorf, 1984) and dose-response evaluation (Howard, Kopta, Krause, & Orlinsky, 1986) are scientific attempts to close the gap between research and practice, providing data that could (and should) be central in policy and practice decisions.

This article, staying within the scientific paradigm mentioned, presents a review of pertinent research, elucidating what the dose-response and randomized clinical trial literatures identify as the optimal treatment "dosage" to facilitate meaningful change. Given the described gap between research and practice, this article then compares these findings on appropriate dosing of treatment with naturalistic data to provide a critical look at how research findings are being utilized in actual clinical practice. The following review has been divided into four sections. The first section discusses clinical significance to define "meaningful change" in treatment. The second section provides a comprehensive review of the dose-response literature, identifying the relationship between treatment duration (or dosage) and treatment outcome (or response). The third section provides an illustrative summary of the psychotherapy efficacy literature, demonstrating expected treatment outcomes under optimal treatment conditions. Finally, the fourth section contrasts these findings with naturalistic data from a large national database of over 6,000 patients receiving treatment within a variety of settings. This comparison illustrates that in many typical treatment delivery systems, patient exposure to treatment and the resulting response fail to meet the standards established by psychotherapy research. Finally, implications for practice, policy, and future research are explored.

CLINICAL SIGNIFICANCE

Whereas statistical significance aims at identifying real differences between samples (or differences within samples at different points in time), it does not indicate the importance of any observed differences. Other methods must be used to specify the meaningfulness of observed differences. In psychotherapy research, one of the most widely used methods for clarifying the clinical meaningfulness of change is clinical significance (Jacobson et al., 1984; Jacobson & Truax, 1991).

The methods originally proposed for clinical significance included two components that must be satisfied before change can be considered clinically meaningful. First, a patient must cross a cutoff point that differentiates a functional population from a dysfunctional one. Thus, the level of functioning of a patient after therapy must be more similar to a functional group than to a patient group before change can be considered meaningful. Second, the change observed must meet a statistical significance criterion by being reliable, that is, greater than the measurement error of the outcome instrument used. It is possible, therefore, for observed change to be statistically significant but not clinically significant. It is not possible, however, to have clinically significant change that is not statistically significant. Clinical significance methods have been modified and elaborated by other researchers to broaden their applicability and utility (Christensen & Mendoza, 1986; Lunnen & Ogles, 1998; Tingey, Lambert, Burlingame, & Hansen, 1996), although these modifications have not always stayed faithful to the "gold standard" of recovery proposed by the original authors.

Clinical significance methodology has received several criticisms, including (1) that it is too stringent a criterion to make it practical, as it is difficult for patients to reach criteria for recovery in the best of circumstances; (2) the improbability, if not impossibility, of some patients with chronic conditions to meet these criteria; and (3) the impossibility for less disturbed patients to meet criteria for clinical significance, as they begin treatment already within the functional range. The extensions to clinical significance methodology proposed by Tingey et al. (1996) were made in part to address these limitations. Despite criticisms, the strengths of clinical significance are many, including its relevance for studies with both large numbers as well as single-subject designs and its ability to facilitate meaningful comparisons between studies. Due to its ability to clearly operationalize meaningful change within groups and at the level of the individual patient, and its applicability to routine clinical practice, clinical significance methodology has much to recommend it as a research and clinical tool. Throughout the review of the psychotherapy outcome literature that follows, clinical significance methodology is used as the primary criterion to distinguish a positive response to treatment.

DOSE-RESPONSE RESEARCH

An area of inquiry that has important practice and policy implications is research on the dose-response relationship in psychotherapy. Dose-response methodology was origi-

nated in the biological sciences for use with biological assays (the study of the potency of stimuli with living subjects). Dose-response research, then, considers the impact of varying doses of a stimulus on a target response variable, such as the impact of differing levels (doses) of insecticide on insect mortality (response). Generally, the sought-after result is the dose-level for a 50% response (Finney, 1971), though probabilities could be estimated for any level (with decreasing accuracy at the two tails of the distribution). Dose-response methods have been widely adopted in medical research and have been adapted to psychotherapy research (Howard et al., 1986).

In psychotherapy research, a "dose" is generally defined as a session of therapy. A "response" is most often defined as whether a particular outcome event (e.g., clinically significant change) has taken place, as measured by change on one or more outcome measures (Howard et al., 1986). Given the binary nature of this definition of response, probit analysis has been traditionally used in dose-response research. This statistical approach is based on probability theory (hence the term probit, or "probability unit"; Hewlett & Plackett, 1979) and includes determining the observed rates and resulting probability of a response for a particular level, or dose, of some treatment (be it a drug, vitamin, poison, or antibiotic). More recently, however, it has been suggested that treatment outcome be observed on a session-by-session basis, rather than depending on pre-post assessments (Anderson & Lambert, 2001; Kadera, Lambert, & Andrews, 1996). This allows the point of change to be directly assessed rather than being based on linear interpolation. This is an important advantage in assessing change, as it has been demonstrated that change often occurs in spurts rather than evenly distributed increments across treatment (Tang & DeRubeis, 1999). Further, when change is monitored session by session, actuarial tables and survival analysis can be used to analyze the resulting longitudinal data. As these techniques directly assess the time, or number of treatment sessions, needed to produce a particular probability of achieving meaningful change, a more accurate estimate of the probability of change occurring at each time point is provided than when relying on pre-post data.

At the heart of estimating the dose-response relationship in psychotherapy is the desire to answer the question: "How much therapy is enough?" Humanistic psychologists inspired by the research efforts of Carl Rogers originally addressed this question. Early efforts of psychotherapy outcome studies concluded that there was a positive relationship between the length of time a patient spent in therapy and the quality of change the patient experienced (Cartwright, 1955; Johnson, 1965; Seeman, 1954; Standal & van der Veen, 1957). Echoing these findings, later research also demonstrated a link between the length of therapy and treatment outcome (Orlinsky & Howard, 1978; Steenbarger, 1994; Strassberg, Anchor, Cunningham, & Elkins, 1977; Weitz et al., 1975). In fact, in a review of 156 papers published on this topic between 1950 and 1992, Orlinsky, Grawe, and Parks (1994) found that 100 (or 64%) studies showed a positive relationship between treatment length and outcome, 50 studies were unable to detect a statistically significant relationship between treatment length and outcome, and only 6 studies found a negative relationship.

Interest in this topic has been growing following the work of Howard et al. (1986). These researchers applied dose-response methodology frequently used in medical/pharmacological research to the question of psychotherapy outcome. They defined "dose" as a session of therapy and "response" as the measured change on a standardized outcome instrument. The authors then proceeded to perform a meta-analytic study using 15 data sets that provided patient outcome as a function of treatment duration. By conducting a probit analysis, they found that between 10% and 18% of the patients had improved prior to the first session of therapy—possibly as a result of the mobilization of resources that entering therapy requires. After 2, 8, and 26 sessions of therapy, 30%, 53%, and 74% of the patients, respectively, had demonstrated improvement. After a year of weekly therapy, 83% of treated patients had improved. One limitation of this initial look at the dose-response relationship in psychotherapy was the definition of improvement that these researchers had available. The summarized studies these authors used reported improvement predominantly via clinician ratings at posttreatment, with a few including patient ratings and others consisting of chart review by researchers. Though the reliability of these ratings of improvement is questionable, this study is important in introducing a powerful methodology to the question of how much therapy is needed for patients to improve.

Using a probit analysis in a similar fashion, McNeilly and Howard (1991) reevaluated Eysenck's (1952) conclusion that psychotherapy is no more effective than spontaneous remission. By reanalyzing Eysenck's data in this

manner, they demonstrated that psychotherapy produced the same effects in 15 sessions as was produced over 2 years by spontaneous remission. This study refutes Eysenck's claim that psychotherapy is ineffective and demonstrates the superiority of an active intervention such as psychotherapy over the simple passage of time in accelerating recovery and reducing patient distress.

Kopta, Howard, Lowry, and Beutler (1994) extended the Howard et al. (1986) study by comparing the differential response rates of symptom types to therapy doses. In this study, 854 outpatients in psychotherapy were administered symptom checklists mainly at the beginning and end of treatment. Results showed that different symptoms improved at different rates, with "acute" symptoms requiring 5 sessions, "chronic" symptoms requiring 14 sessions, and "characterological" symptoms requiring 104 sessions for a 50% response. This was the first published study to utilize clinical significance methodology in a dose-response analysis. In a similar study, Maling, Gurtman, and Howard (1995) compared the rate of improvement of interpersonal problems in psychotherapy. They found that interpersonal difficulties, like psychiatric symptoms, respond to treatment at different rates. For instance, problems with control responded rapidly, with a 50% response rate occurring within 10 sessions, whereas problems with self-effacement showed almost no response to psychotherapy even after many sessions. Rate of improvement for social detachment problems fell in the middle range, with a 30% response after 17 sessions and a 55% response after 38 sessions. Similarly, Barkham et al. (1996) found that 50% of depressed patients, whether receiving psychodynamic-interpersonal or cognitive-behavioral treatment, reached clinically significant change in psychiatric symptoms after 8 sessions of therapy, whereas it took 16 sessions to achieve a 40% rate of clinically significant improvement in interpersonal difficulties.

Kadera et al. (1996) presented a new perspective on dose-response research with their study of 64 adult outpatients. This study followed each patient on a session-by-session basis for a 10-month period of time. Rather than using a probit analysis as Howard et al. (1986) did, the data were analyzed using an actuarial approach based on life tables. Using these methods, Kadera et al. found that patients improved at a slower rate than was reported by Howard and his colleagues. The Kadera et al. study represents a methodological evolution that was able to address a shortcoming in earlier dose-response methodology—namely, its reliance on pre-post testing to estimate session-by-session change. Current trends in dose-response methodology have built on the Kadera et al. procedures by using survival analysis (Singer & Willett, 1991), along with clinical significance methodology to provide a direct session-by-session analysis of change. For instance, Anderson and Lambert (2001) found that, on average, a "dosage" of 13 sessions of psychotherapy was needed to reach a modest 50% improvement rate. Unfortunately, generalization from the Kadera et al. and the Anderson and Lambert studies is limited due to small numbers and the fact that therapy was delivered by graduate student trainees rather than licensed professionals. To overcome these limitations, Hansen and Lambert (in press) attempted to replicate the Kadera et al. and Anderson and Lambert studies with a larger number of patients drawn from a wider variety of treatment settings. The study provided a naturalistic look at psychotherapy within "routine practice" and offered a view of how current treatment practices met patient needs. While there were significant differences observed in rates of improvement across these sites, between 15 and 20 sessions of therapy were typically needed to observe a 50% rate of recovery among patients receiving treatment, regardless of where they were being treated.

Table 1 provides a summary of the dose-response literature and gives an overview of what researchers have found in attempting to address the question of how much therapy is enough. A convention of dose-response research, which has been followed here, has been to report the median survival or response time, that is, the point at which 50% of a study's subjects have demonstrated a positive response to treatment. It can be argued that selecting the median response point as a benchmark is problematic as it may be a high mark for some disorders (i.e., Axis II and schizophrenia), and a low mark for others (i.e., phobias, panic disorder). For the current review, the 50% response value was maintained as the benchmark, however, as few of the studies summarized in this section presented detailed data on the diagnostic groups treated or the treatments offered. The studies reviewed tended to represent a wide range of disorders, although the majority were Axis I mood and anxiety disorders. Due to the heterogeneous disorders and treatments represented here, selecting a more refined response value is unwarranted. And, despite potential limitations, the median value does provide a useful measurement—the point at which, on average, half of those seeking treatment respond positively to treatment.

Table 1. Summary of dose-response study findings on number of sessions required to reach a 50% patient improvement rate

Reference	Number of Sessions	Comments
Howard et al. (1986)	8	Did not use clinical significance, only pre-post comparisons
Kopta et al. (1994)	5, 14, 104	Session numbers refer to 50% response in acute, chronic, and characterological symptoms, respectively
Maling, Gurtman, & Howard (1995)	10, 38	Session numbers refer to 50% response in problems with control and social detachment, respectively
Barkham et al. (1996)	8, 16+	Eight sessions for 50% symptom improvement, 16 sessions for 40% interpersonal problem improvement
Kadera, Lambert, & Andrews (1996)	16	Followed patients session by session, found flatter rate of improvement than the previous pre-post designs
Anderson & Lambert (2001)	13	Used survival analysis on patient data collected from each session
Hansen & Lambert (in press)	18	Average survival time computed from treatment sites in study

In this review, a positive response to treatment is typically defined as a patient exhibiting clinically significant change. One problem with comparing values in Table 1, however, is that not all studies have used the same criteria for evaluating change. Some studies focused only on psychiatric symptoms, whereas some included an evaluation of broader symptom types. Some used clinical significance methods, whereas others did not. Some studies were naturalistic and included a wide range of patients and diagnoses, as well as treatment types, whereas other studies were carefully controlled clinical trials with carefully screened patients and closely monitored treatments. Despite these differences, however, there is a surprising consensus across a number of these studies, and it can be argued, in the broadest sense, that some number of sessions greater than 10 but fewer than 20 is typically required before 50% of patients meet criteria of recovery. A realistic summary of this literature suggests that between 13 and 18 sessions of therapy are needed for psychiatric symptom alleviation, across various types of treatment and patient diagnosis.

OUTCOME RATES IN RANDOMIZED CLINICAL TRIALS RESEARCH

Randomized clinical trials are essential in demonstrating the efficacy of specific treatments beyond no-treatment conditions, minimal-treatment control conditions, and even competing-treatment conditions. Clinical trial research has been criticized, however, for trading external validity (making research generalizable) for internal validity (making research replicable). Some even question the ability of clinical trials to inform practice in a meaningful way (Howard et al., 1996; Seligman, 1996; Wampold, 1997). As clinical trials are aimed at identifying treatment effects and making causal conclusions, careful design and control of variables are required. Nathan et al. (2000) identify four areas of requiring special attention for clinical trials: (1) appropriate control conditions, (2) random assignment of subjects, (3) treatment manuals, and (4) well-defined patient groups. Typically, clinical trials apply manualized and highly managed treatment, delivered by highly trained clinicians, to carefully selected but randomly assigned patient samples under optimal treatment conditions. The validity of studies carried out under these conditions is questionable, as these conditions are not typically found together in routine clinical practice. This has led to a call for studies on psychotherapy effectiveness, or even studies integrating efficacy and effectiveness, addressing the utility and potency of interventions in real world settings (Howard et al., 1996; Klein & Smith, 1999; Kopta et al., 1999; Norquist, Lebowitz, & Hyman, 1999; Seligman, 1995, 1996).

Despite these concerns about the generalizability of clinical trials to clinical practice, studies applying "empirically supported treatments" to real world settings have found outcomes similar to those in clinical trials (Franklin, Abramowitz, Kozak, Levitt, & Foa, 2000; Stuart, Treat, & Wade, 2000; Wade, Treat, & Stuart, 1998). Shadish and colleagues have also conducted meta-analyses attempting to control for numerous variables, including the representativeness of clinical settings (Shadish et al., 1997, 2000). These authors failed to find a significant effect for outcome based on where a study falls on the continuum of efficacy/effectiveness.

In light of the debate between efficacy and effectiveness, it seems useful to review the treatment outcomes reported in clinical trials to attempt to understand what can be achieved through an optimal application of psychotherapy. Table 2 summarizes a representative sample of randomized clinical trial studies from the psychotherapy

Table 2. A representative sample of treatment durations and outcomes in randomized clinical trials

Reference	N	# Sess.	Treatment Response	Treatment Type[a]	Population
Arntz & Van Den Hout (1996)	36 subjects, 3 conditions	12	78% of CBT and 50% of BT subjects panic free	CBT, BT	Panic disorder
Barkham et al. (1999)	116 subjects, 2 conditions	3	No difference between conditions, 65% to 72%, depending of initial severity of symptoms	CBT, PI	"Subsyndromal" depression
Barlow et al. (1989)	56 subjects, 4 conditions	15	Panic free: BT = 40%, CBT = 79%, both = 74%	BT, CBT, both	Panic disorder
Barlow et al. (2000)	312 subjects,	11	Clinically significant response = 48.7% CBT, 45.8% IM, CBT + IM = 60.3%, CBT + PLA = 57.1%, PLA = 21.7%	CBT, IM, CBT + IM, CBT + PLA	Panic disorder
Barlow, Rapee, & Brown (1992)	65 subjects, 4 conditions	15	Treatment response: BT = 36%, CBT = 50%, both = 33%	BT, CBT, both	Generalized anxiety disorder
Beck et al. (1994)	64 subjects, 3 conditions	10	82% CBT and 68% AMT subjects improved; 65% of CBT and 47% of AMT subjects panic free	CBT, AMT	Panic disorder
Black et al. (1993)	75 subjects, 3 conditions	8	81% of FL, 53% of CBT subjects panic free	FL, CBT	Panic disorder
Blackburn & Moore (1997)	75 subjects, 3 conditions	16	CBT conditions: 24%, 24% recovered; 67%, 65% recovered or improved; ADM condition: 35% recovered, 68% recovered or improved	CBT, ADM	Major depressive disorder
Borkovec & Costello (1993)	55 subjects, 3 conditions	12	Response across a variety of measures was 33.3% for NDT, 83.3% for AMT, and 78.9% for CBT	NDT, AMT, CBR	Generalized anxiety disorder
Bouchard et al. (1996)	28 subjects, 2 conditions	15	Percent panic free vs. high end state functioning: BT = 79% vs. 86%, CBT = 64% vs. 64%	BT, CBT	Panic disorder
Butler et al. (1991)	57 subjects, 3 conditions	4–12	Recovery across three meaures is 32% for CBT and 16% for BT	CBT, BT	Generalized anxiety disorder
Clark et al. (1994)	64 subjects, 4 conditions	12–15	Percent panic free versus high end state functioning: CBT = 90% vs. 80%, AMT = 50% vs. 25%, IM = 55% vs. 40%	CBT, AMT, IM	Panic disorder
Clark et al. (1999)	43 subjects, 3 conditions	Full CBT = 12, brief CBT = 5	No difference between conditions, panic free versus high end state for both full and brief CBT = 79% vs .71%	Full CBT, brief CBT	Panic disorder
Durham et al. (1994)	80 subjects, 3 conditions	8–20, 16–	No difference due to treatment length, percent recovered: CBT = 68%, AMT = 44%, PI = 34%	CBT, PI, AMT	Generalized anxiety disorder
Foa et al. (1991)	45 subjects, 4 conditions	9	Recovery = 71% for CBT, 40% for BT, 18% for NDT	CBT, BT, NDT	PTSD
Freeston et al. (1997)	29 subjects, 2 conditions	Mean CBT = 25.7	77% of subjects had clinically significant change at posttreatment	CBT	Obsessive compulsive disorder
Hollon et al. (1992)	107 subjects, 4 conditions	12	Treatment response on BDI: CBT = 44%, IM = 40%, both = 48%	CBT, IM, both	Major depressive disorder
Jacobson et al. (1996)	150 subjects, 3 conditions	12–20	51.5% recovered, 62.3% improved or recovered	CBT	Major depressive disorder
Klosko et al. (1990)	57 subjects, 4 conditions	15	% panic free: BT = 87%, AZ = 50%, PLA = 36%	BT, AZ, PLA	Panic disorder
Ladouceur et al. (2000)	26 subjects, 2 conditions	16	77% of subjects do not meet diagnostic criteria after treatment	CBT	Generalized anxiety disorder
Murphy et al. (1984)	87 subjects, 4 conditions	12	No difference between conditions, 63% not depressed after treatment	CBT, ADM, CBT + PLA, CBT + ADM	Major depressive disorder
Nezu (1986)	26 subjects, 3 conditions	8	Reliable change on BDI: CBT group = 90.9%, NDT group = 22.2%	Group CBT, group NDT	Major depressive disorder
Nezu & Perri (1989)	39 subjects, 3 conditions	10	Recovery = 85.7% for full treatment, 50% for partial treatment	Group full CBT, group partial CBT	Major depressive disorder
Ogles, Lambert, & Sawyer (1995)	162 subjects, 4 conditions	12–15	Recovery = 65% for CBT, 85% for PI, 82% for IM	CBT, PI, IM	Major depressive disorder
Shear et al. (1994)	45 subjects, 2 conditions	15	No difference between conditions, 78% of NDT and 66% of CBT subjects panic free	NDT, CBT	Panic disorder
Thompson, Gallagher, & Breckenridge (1987)	91 subjects, 3 conditions	16–20	No difference between conditions, 70% of sample no longer depressed, 50% response on BDI, 75% response on HRSD	CBT, BT, PI	Major depressive disorder
van Oppen et al. (1995)	71 subjects, 2 conditions	16	% recovered vs. recovered or improved: CBT = 50% vs. 75%, BT = 28% vs. 66%	CBT, BT	Obsessive compulsive disorder
Williams & Falbo (1996)	48 subjects, 4 conditions	8	No difference between conditions, % panic free: low agoraphobia = 94%, high agoraphobia = 52%	CBT, BT, both	Panic disorder

[a]Although many different treatment are summarized here, they have been collapsed into the somewhat artificial categories below: ADM = Anti-Depressant Medication; AMT = Anxiety Management Training; AZ = Alprozalom; BT = Behavior Therapy; CBT = Cognitive-Behavior Therapy; FEP = Focused Expressive Therapy; FL = Fluvoxamine; HT = Hypnotherapy; IM = Imipramine; NDT = Nondirective Therapy; PI = Psychodynamic/Interpersonal Therapy; PLA = Pill Placebo.

research literature. The studies reflected in Table 2 were selected to be diverse in treatment duration, type of therapy, and type of disorder treated. While this diversity may limit the generalizability of the findings presented here to any specific context, the goal is to present a broad overview of what can be expected in the optimal application of psychotherapy. This is not an attempt to provide an exhaustive review of the literature, as other reviews have done a more thorough job in identifying and reviewing the pertinent literature for specific contexts (e.g., Abramowitz, 1997; DeRubeis & Crits-Cristoph, 1998; Durham & Allan, 1993; Nietzel, Russell, Hemmings, & Gretter, 1987; Roth & Fonagy, 1996; van Balkom et al., 1997).

A number of search strategies were utilized in selecting articles for this review. First, meta-analyses and review articles focusing on psychotherapy outcome research and empirically supported treatments were identified (Abramowitz, 1997; Anderson & Lambert, 1995; Crits-Cristoph, 1992; DeRubeis & Crits-Cristoph, 1998; Durham & Allan, 1993; Gaffan, Tsaousis, & Kemp-Wheeler, 1995; Gould, Buckminster, Pollack, Otto, & Yap, 1997; Gould, Otto, Pollack, & Liang, 1997; Gould, Otto, & Pollack, 1995; Lambert & Bergin, 1994; Nietzal et al., 1987; Roth & Fonagy, 1996; Shadish et al., 1997, 2000; Stevens, Hynan, & Allen, 2000; Taylor, 1996; van Balkom et al., 1997; Wampold et al., 1997), and studies reviewed and referenced in these articles were scrutinized for inclusion. Second, the indices of major psychiatry and psychology journals from 1980 to present were examined (including *American Journal of Psychiatry, Archives of General Psychiatry, Behavior Therapy, Behaviour Research and Therapy, Journal of Consulting and Clinical Psychology,* and *Journal of Nervous and Mental Disease*). Third, a computer search was conducted using the Psych INFO database from 1967 to February 2001. Finally, the reference lists of articles identified during this search were also reviewed for potential studies to be included.

The criteria used in selecting studies for inclusion in Table 2 from potential studies identified above included (1) the experimental condition must be a legitimate psychotherapeutic treatment (examples include behavioral therapy, cognitive-behavioral therapy, psychodynamic therapy, interpersonal therapy, client-centered therapy, gestalt therapy, and hypnotherapy); (2) studies must use random assignment of subjects into either the experimental or control conditions; (3) comparison conditions must include waitlist controls, minimal treatment groups, medication treatment conditions, medication placebo conditions, or alternative psychotherapeutic treatments; (4) treatment must be delivered by either experienced therapists or therapists who have received special training in the treatment conditions, with treatment manuals and supervision provided; (5) therapists must be experienced in delivering the experimental treatment or there must be observation, supervision, or other procedure to ensure treatment adheres to protocol; (6) treatment outcomes must be expressed using clinical significance criteria to determine percentage of subjects responding to treatment; and (7) treatment must target adult patients with nonpsychotic Axis I disorders. The last criterion was selected to make the treatment samples in the reviewed literature correspond more closely with the dose-response data reviewed previously and the naturalistic data presented in the next section. Selecting studies utilizing clinical significance criteria limits the studies that can be included to those that were published from the mid-1980s to the present, as clinical significance methodology was introduced in 1984.

Treatment response was defined differently within the studies reviewed in this section, as well as in the measures of change used. Also, Table 2 presents a wide range of diagnoses, which may respond to treatment at different rates and through different mechanisms. Further, though a majority of the reviewed studies feature behavioral or cognitive-behavioral therapies, a wide variety of treatments may differ in active change components and therefore in observed patterns of change over time. Therefore, comparisons between these studies should be made cautiously. However, for this review, the focus is on treatment in general; therefore, the studies summarized in Table 2 are being pooled to produce a summary of treatment effects in clinical trials research. This summary should be viewed as a rough estimate of treatment expectations when experienced or highly trained therapists provide a manualized and specifically targeted treatment for a specific disorder or problem under fairly optimal treatment conditions, which include supervision or special attention given to treatment delivery.

The 28 studies summarized in Table 2 represent 2,109 patients in 89 treatment conditions. For this summary, only treatment conditions that represent an active psychotherapeutic treatment were included. Wait-list control groups, minimal treatment conditions, and pill-placebo conditions were excluded from the analysis. It is interesting to note the similarity in recommended treatment length from

the reviewed dose-response research (13 to 18) sessions and the length of treatments provided in typical clinical trial research. The average number of sessions received across these treatment conditions was 12.7 (SD = 4.6). Although there was variability in the lengths of treatment provided in the summarized literature, there was not a significant relationship between treatment length and outcome across the studies summarized in Table 2. The rate of response to treatment using clinical significance criteria was 57.6%. Using the less conservative measure of reliable change, the response rate was 67.2%. While a few of these values are based on patients completing treatment, and may be reduced somewhat if therapy dropouts are considered, the aggregate results suggest that on average in clinical trials research, more than half of those entering treatment reach recovery, and about two-thirds have meaningful improvement. These response rates are slightly higher than those predicted within 13 to 18 sessions by the dose-response literature cited previously.

TREATMENT DURATION AND OUTCOME IN ROUTINE PRACTICE

Given the information presented in the previous two sections (summarized in Tables 1 and 2), a logical next step is to examine a large number of patients drawn from routine practice across a variety of treatment settings to determine how well the summarized dose-response and clinical trial literatures capture typical treatment delivery. This allows two critical questions to be addressed: Are patients in routine practice being exposed to sufficient amounts of psychotherapy? Are patients in routine practice responding to psychotherapy as expected? For this review, data were obtained from a number of archival data sources. The data were selected to reflect a naturalistic perspective of psychotherapy across a variety of settings.

A total of 6,072 patients across six data collection sites were utilized. Notably, a large number of patients (3,101, or 33% of the original sample) were excluded from this data set because they received only one session of therapy (patients with only one session of therapy were excluded as one data point does not provide the required information to compute a change score). This finding corresponds with the report of Garfield (1994), who notes that around 25% to 50% of patients within numerous studies and across diverse treatment settings "refuse psychotherapy" by failing to return to treatment after an initial intake or therapy session. In fact, in this review, some data collection sites provided only data on subjects with repeated measurements, suggesting that even greater than 33% of patients in naturalistic settings receive only one session of therapy. The data sites, number of subjects from each site, and the average number of sessions for patients at each site are listed in Table 3.

Table 3. Data sample sources, sample size, and average number of sessions

Sample Source	Sample Size	M (SD) # Sessions	Median # Sessions
Employee Assistance Program	3,269	3.6 (2.0)	3.0
University Counseling Center	1,188	5.8 (5.4)	4.0
Local HMO	595	3.3 (2.4)	2.0
National HMO	536	5.1 (4.0)	4.0
Training CMH	123	9.5 (6.8)	8.0
State CMH	361	4.1 (2.8)	4.0
Total	6,072	4.3 (3.5)	3.0

The largest number of participants (n = 3,269) were enrolled in a nationwide employee assistance program (EAP) offering short-term treatment for adjustment problems. In addition, 1,188 participants received care at a university counseling center based at a large Western university. A third category of participants included members of local (n = 595) and national (n = 536) health maintenance organizations (HMOs). A further source providing participants was a university-supported training clinic in a Western state (n = 125). This site serves low-income and underinsured populations who are experiencing moderate levels of distress and life problems. The remaining participants (n = 361) were drawn from a state-supported community mental health service (CMH) within a large Northern state.

Each patient included in the data set received psychological treatment scheduled on a weekly basis and completed the Outcome Questionnaire (OQ-45; Lambert et al., 1996a, 1996b) prior to each session of therapy, including the first (i.e., at intake). The OQ-45 is a 45-item outcome measure designed to assess patient functioning across three domains: Intrapsychic distress, interpersonal problems, and social role functioning. The OQ-45 has demonstrated adequate psychometric properties (internal consistency α = .93; test-retest reliability = .84; correlations ranging from .53 to .88 with similar instruments; Lambert et al., 1996a, 1996b; Umphress, Lambert, Smart, Barlow, & Clouse, 1996) and sensitivity to change (Vermeersch, Lambert, & Burlingame, 2000).

After collecting and compiling the OQ-45 data from the data sites, we used the resulting data for each individual patient to determine the clinical significance of the patient's change in treatment. The clinical significance methods utilized here included using the cutoff score (a score of 63 on the OQ-45) to distinguish between functional and dysfunctional groups, as determined by normative data established by Lambert et al. (1996a). Further, the RCI (a change of 14 points on the OQ-45) was used to define four categories of therapy outcome, including (1) *deteriorated,* meaning a patient's OQ-45 score had reliably moved in a negative direction during the course of therapy as judged by the RCI value; (2) *no change,* meaning a patient's OQ-45 score had not changed reliably in any direction over the course of therapy; (3) *improved,* meaning a patient's OQ-45 score had reliably changed in a positive direction over the course of therapy as judged by the RCI value; and (4) *recovered,* meaning a patient's OQ-45 score had improved reliably as judged by the RCI, as well as having moved from within the range of the dysfunctional distribution to within the range of the functional distribution during the course of therapy. Table 4 illustrates the observed outcomes on the OQ-45 by site.

Table 5 presents a sobering picture of routine clinical practice. This table illustrates the median number of sessions that patients received at the treatment sites in this summary. This information, though disappointing, is consistent with Garfield's (1994) review of literature published from the 1940s to 1989 showing that the median number of sessions reported by studies was typically between 4 and 10. The more distressing information contained in Table 5 is the low percentages of patients who met clinical significance criteria for recovery after the median number of sessions. The site exhibiting the most successful patients after this brief exposure to psychotherapy had fewer than 10% of its patients recover. When criteria were relaxed to include those patients who demonstrated meaningful improvement (whether or not they "recovered"), the response rate was still less than 25%. While much of the data summarized here was collected from sites influenced by managed health care organizations, there is little in the literature to suggest that the findings presented here are not representative of practice in general. Whether this is true cannot be known until effectiveness research is produced that demonstrates otherwise—that is, that private practitioners and mental health providers not directly influenced by managed health care organizations produce more adequate treatment dosing with improved response rates.

Table 4. Number of patients, by site, who demonstrated reliable change

Site	Deteriorated	No Change	Improved	Recovered
Employee Assistance Program	216 (6.6%)	1,911 (58.5%)	645 (19.7%)	497 (15.2%)
University Counseling Center	115 (9.7%)	684 (57.6%)	239 (20.1%)	150 (12.6%)
Local HMO	84 (14.1%)	321 (53.9%)	122 (20.5%)	68 (11.4%)
National HMO	40 (7.5%)	258 (48.1%)	153 (28.5%)	85 (15.9%)
Training CMH	4 (3.2%)	57 (45.6%)	39 (31.2%)	25 (20.0%)
State CMH	37 (10.2%)	219 (60.7%)	74 (20.5%)	31 (8.6%)
Total	496 (8.2%)	3,448 (56.8%)	1,272 (20.9%)	856 (14.1%)

Table 5. Percentage of patients, by site, who achieve clinically meaningful improvement within median treatment length

Site	Sample Size	Median # Sess.	% Who Recover	% Who Improve
Employee Assistance Program	3,269	3	7.4%	18.3%
University Counseling Center	1,188	4	5.9%	15.2%
Local HMO	595	2	5.7%	14.3%
National HMO	536	4	9.1%	24.4%
Training CMH	123	8	6.5%	20.3%
State CMH	361	4	5.8%	17.7%
Total	6,072	3	6.5%	16.6%

Combining conclusions gleaned from Tables 1 and 5, we can see that half of the patients receiving psychotherapy receive only about a quarter of the length of treatment that the literature has noted as necessary to observe a 50% response rate. Given the disappointing response rates of patients after the median number of sessions within each site, and comparing them with the estimates provided in Table 2, we believe the response rates observed in routine practice fell well short of what may be expected, given the research on clinical trials reviewed here.

Looking to Table 4 again, we can see that the patients who continue in treatment, after completing the median number of sessions, continue to show improvement, which is what would be hoped and expected, that is, that providing more treatment results in greater patient response. Interestingly, the site that ultimately had the best response rate, the training CMH, is also the site that had

the most sessions delivered per patient and the highest median treatment length.

CONCLUSIONS AND IMPLICATIONS FOR PRACTICE

It is useful to consider this review's findings in relation to the research summarized, noting that the manner in which the dose-response relationship was operationalized varied between these studies, and therefore resulted in considerable variability in results. The studies that are perhaps the most meaningful (due to improved technical application of methodological refinements, including the use of psychometrically sound outcome instruments, clinical significance methodology, and session-by-session monitoring of outcome) are the Kadera et al. (1996), Anderson and Lambert (2001), and Hansen and Lambert (in press) studies. These studies are fairly consistent in indicating the amount of therapy needed, and the rate of the corresponding improvement (which typically occurred at a slower rate and had a flatter response curve than that found by Howard et al., 1986).

One of the most striking findings in this report is that most patients did not receive enough psychotherapy to reach even a moderate level of clinically meaningful change. As shown in Table 3, the mean and median number of treatment sessions delivered to each subject was extremely low. At all sites but one, the mean was between three and five sessions, and the median was between two and four sessions—and this was calculated after all patients with only one session of therapy had been dropped from the study. These values are far below the number of sessions assumed necessary in clinical trials to produce improvement, let alone recovery. Even a very modest level of positive outcome, such as 50% improvement, would still require as many as 18 sessions of therapy for the average patient to reach this level. It is a rare patient, however, who is actually exposed to 18 sessions of therapy.

As previously noted, the psychotherapy outcome literature indicates a correlation between the number of sessions of therapy received and the amount of improvement within patients (Orlinsky et al., 1994; Steenbarger, 1994). Previous studies suggest that this improvement does level off and reach a point of diminishing returns over time (Howard et al., 1986), although this also depends on the outcome domain measured. Some domains, such as interpersonal problems, continue to show improvement after other domains, such as symptom distress, have reached an apparent floor effect (Barkham et al., 1996; Hansen, Umphress, & Lambert, 1998; Kopta et al., 1994).

A reasonable conclusion based on this review, then, is that more treatment (i.e., longer treatment) is better, and the treatment typically received by patients in naturalistic settings is insufficient by most standards of care. Of course, continued research is needed to further explore the amount of treatment sufficient for given patients in particular contexts. For instance, the dose-response research summarized here suggests that there is a limit at which point further treatment results in diminishing gains, especially for psychiatric symptoms. Much more work should be done to define acceptable treatment guidelines, but in the meantime, treatment guidelines that emphasize *time to clinically significant change* need to be established.

This review is not without limitations, the most striking of which is the lack of information about the naturalistic samples. First, little is known about the therapists providing services within the various sites. Some are predominantly master's-level therapists, many with training in social work. Other sites employ predominately PhD- or PsyD-level therapists, while services at the training CMH site are provided by unlicensed trainees in doctoral or master's level training programs. No information is available on years of experience or type of therapy practiced. Second, little is known about the patients composing these samples. For instance, there is no information available on patient diagnosis, age, ethnicity, education, or previous treatment history. Third, there is no information on psychiatric treatments—whether patients have used medication previously or within the studied treatment episode. Fourth, no follow-up information is available on these patients. Fifth, the reasons for patient discontinuation of therapy across sites are unknown. While we can surmise that to some degree differences between observed retention rates when comparing clinical trials and naturalistic settings can be attributed to incentives (participants in studies generally have incentives to participate, while patients in treatment—and to some degree therapists as well—have incentives to end treatment), much more needs to be known about why patients choose to remain in or leave treatment. Finally, the bulk of the data is from insurance-driven treatments, representing patients who have stable employment, and a large subset is from a university counseling center. While the training CMH and the state CMH serve low SES patients, and the state CMH serves more chronic mental illness and substance abuse

patients than the other sites, these populations were underrepresented in these data. Given these limitations, this review should be considered as a general overview and perhaps a warning to the field, but it lacks the specificity to allow inference in specific instances.

The many implications of this review can be classified into three broad categories: (1) implications for clinical practice, (2) implications for mental health policy, and (3) implications for clinical training. First, in clinical practice, delivering service that has adequate duration is an important concern. Holding patients in treatment until satisfactory gains are achieved is a challenging task at times and should be a focus of clinical practice. But it requires clinicians to adopt a formal system of measuring and monitoring client treatment response in relation to normative standards. Lambert et al. (2001) demonstrated that this could be achieved and that feedback to therapists on patient progress resulted in nearly doubling the amount of therapy they received and significant gains in treatment outcome. In addition, clinicians might consider the advantages of educating or orienting patients to treatment with regard to the expected duration of a complete course of therapy.

Implications for mental health policy are also many and are at odds with some of the practices currently utilized. For instance, rather than requiring justification for continuing an ongoing treatment, mental health managers may do better to require justification for termination prior to completing a course of treatment—with justification for ongoing treatment required only after a typical treatment course has been delivered. Further, it would be more efficient for mental health management to focus on cases identified as not improving according to expectations, rather than managing all cases in a practice (Lambert et al., 2001). Finally, at the very least, treatment limits should be set to reflect the duration of treatment needed to achieve adequate treatment gains—according to this review, this would be well beyond 20 sessions if more than 50% of patients are to experience a clinically significant gain.

This review also has implications for clinical training. First, it appears to be important for clinicians in training to be able to demonstrate competence in conducting at least one method of empirically supported treatment, if not several. Further, trainees should be aware of expected treatment durations and the gains that can be achieved, as well as how to monitor their own treatment. In this way, clinicians can identify patients who do not improve as expected and can alter treatment plans, review diagnoses, get psychiatric consultations, or receive supervision when needed to aid patients in succeeding in treatment. Finally, trainees need to be able to orient and educate patients in the dosage of therapy that may be needed for an adequate treatment response and in methods of enhancing motivation (Garfield, 1994).

Much work remains to be done to ensure that patients receive adequate treatment. Patient-focused research that concentrates on treatment response in relation to "dosage" is of particular importance because clinical trials fail to inform us about routine clinical practice and the degree to which optimally effective treatments are offered. Research that measures patient treatment response after every session is likely to pay big dividends for patient health and serve as a guide for administrators who wish to change their focus from limiting treatment to ensuring that patients receive sufficient services to get the needed benefit.

ACKNOWLEDGMENT

This work was supported, in part, by Grant 5 T32 MH20031 from the National Institute of Mental Health. We thank Angela Hansen for her helpful contributions in the preparation of this article.

REFERENCES

Abramowitz, J. S. (1997). Effectiveness of psychological and pharmacological treatments for obsessive-compulsive disorder: A quantitative review. *Journal of Consulting and Clinical Psychology, 65,* 44–52.

Anderson, E. M., & Lambert, M. J. (1995). Short-term dynamically oriented psychotherapy: A review and meta-analysis. *Clinical Psychology Review, 15,* 503–514.

Anderson, E. M., & Lambert, M. J. (2001). A survival analysis of clinically significant change in outpatient psychotherapy. *Journal of Clinical Psychology, 57,* 875–888.

Arntz, A., & van den Hout, M. (1996). Psychological treatment of panic disorder without agoraphobia: Cognitive therapy versus applied relaxation. *Behaviour Research and Therapy, 34,* 113–121.

Barkham, M., Rees, A., Stiles, W. B., Shapiro, D. A., Hardy, G. E., & Reynolds, S. (1996). Dose-effect relations in time-limited psychotherapy for depression. *Journal of Consulting and Clinical Psychology, 64,* 927–935.

Barkham, M., Shapiro, D. A., Hardy, G. E., & Rees, A. (1999). Psychotherapy in two-plus-one sessions: Outcomes of a randomized controlled trial of cognitive-behavioral and psychodynamic-interpersonal therapy for subsyndromal

depression. *Journal of Consulting and Clinical Psychology, 67,* 201–211.

Barlow, D. H., Craske, M. G., Cerny, J. A., & Klosko, J. S. (1989). Behavioral treatment of panic disorder. *Behavior Therapy, 20,* 261–282.

Barlow, D. H., Gorman, J. M., Shear, K. M., & 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.

Barlow, D. H., Rapee, R. M., & Brown, T. A. (1992). Behavioral treatment of generalized anxiety disorder. *Behavior Therapy, 23,* 551–570.

Beck, J. G., Stanley, M. A., Baldwin, L. E., Deagle, E. A. III, & Averill, P. M. (1994). Comparison of cognitive therapy and relaxation training for panic disorder. *Journal of Consulting and Clinical Psychology, 62,* 818–826.

Black, D. W., Wesner, R., Bowers, W., & Gabel, J. (1993). A comparison of fluvoxamine, cognitive therapy, and placebo in the treatment of panic disorder. *Archives of General Psychiatry, 50,* 44–50.

Blackburn, I. M., & Moore, R. G. (1997). Controlled acute and follow-up trial of cognitive therapy and pharmacotherapy in outpatients with recurrent depression. *British Journal of Psychiatry, 171,* 328–334.

Borkovec, T. D., & Costello, E. (1993). Efficacy of applied relaxation and cognitive-behavioral therapy in the treatment of generalized anxiety disorder. *Journal of Consulting and Clinical Psychology, 61,* 611–619.

Bouchard, S., Gauthier, J., Laberge, B, French, D., Pelletier, M., & Godbout, C. (1996). Exposure versus cognitive restructuring in the treatment of panic disorder with agoraphobia. *Behavior Research and Therapy, 34,* 213–224.

Butler, G., Fennell, M., Robson, P., & Gelder, M. (1991). Comparison of behavior therapy and cognitive behavior therapy in the treatment of generalized anxiety disorder. *Journal of Consulting and Clinical Psychology, 59,* 167–175.

Cartwright, D. S. (1955). Success in psychotherapy as a function of certain actuarial variables. *Journal of Consulting Psychology, 19,* 357–363.

Chambless, D. L., Sanderson, W. C., Shoham, V., Johnson, S. J., Pope, K. S., Crits-Christoph, P., Baker, M., Johnson, B., Woody, S. R., Sue, S., Beutler, L., Williams, D. A., & McCurry, S. (1996). An update on empirically validated therapies. *Clinical Psychologist, 49,* 5–18.

Christensen, L., & Mendoza, J. L. (1986). A method of assessing change in a single subject: An alteration of the RC index. *Behavior Therapy, 17,* 305–308.

Clark, D. M., Salkovskis, P. M., Hackmann, A., Middleton, H., Anastasiades, P., & Gelder, M. (1994). A comparison of cognitive therapy, applied relaxation and imipramine in the treatment of panic disorder. *British Journal of Psychiatry, 164,* 759–769.

Clark, D. M., Salkovskis, P. M., Hackmann, A., Wells, A., Ludgate, J., & Gelder, M. (1999). Brief cognitive therapy for panic disorder: A randomized controlled trial. *Journal of Consulting & Clinical Psychology, 67,* 583–589.

Crits-Christoph, P. (1992). The efficacy of brief dynamic psychotherapy: A meta-analysis. *American Journal of Psychiatry, 149,* 151–158.

DeLeon, P. H. (1988). Public policy and public service. *American Psychologist, 43,* 309–315.

DeLeon, P. H., Vandenbos, G. R., & Bulatao, E. Q. (1991). Managed mental health care: A history of the Federal Policy Initiative. *Professional Psychology: Research and Practice, 22,* 15–25.

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.

Durham, R. C., & Allan, T. (1993). Psychological treatment of generalised anxiety disorder: A review of the clinical significance of results in outcome studies since 1980. *British Journal of Psychiatry, 163,* 19–26.

Durham, R. C., Murphy, T., Allan, T., Richard, K., Treliving. L. R., & Fenton, G. W. (1994). Cognitive therapy, analytic psychotherapy and anxiety management training for generalised anxiety disorder. *British Journal of Psychiatry, 165,* 315–323.

Eysenck, H. J. (1952). The effects of psychotherapy: An evaluation. *Journal of Consulting Psychology, 16,* 319–324.

Finney, D. J. (1971). *Probit analysis.* Cambridge: Cambridge University Press.

Foa, E. B., Rothbaum, B. O., Riggs, D. S., & Murdock, T. B. (1991). Treatment of posttraumatic stress disorder in rape victims: A comparison between cognitive-behavioral procedures and counseling. *Journal of Consulting and Clinical Psychology, 59,* 715–723.

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 compare with nonrandomized samples. *Journal of Consulting and Clinical Psychology, 68,* 594–602.

Freeston, M. H., Ladouceur, R., Gagnon, F., Thibodeau, N., Rheaume, J., Letarte, H., & Bujold, A. (1997). Cognitive-behavioral treatment of obsessive thoughts: A controlled study. *Journal of Consulting and Clinical Psychology, 65,* 405–413.

Gaffan, E. A., Tsaousis, I., & Kemp-Wheeler, S. M. (1995). Researcher allegiance and meta-analysis: The case of cognitive therapy for depression. *Journal of Consulting and Clinical Psychology, 63,* 966–980.

Garfield, S. L. (1994). Research on client variables in psychotherapy. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4th ed., pp. 190–228). New York: Wiley.

Gould, R. A., Buckminster, S., Pollack, M. H., Otto, M. W., & Yap, L. (1997). Cognitive-behavioral and pharmacological treatment for social phobia: A meta-analysis. *Clinical Psychology: Science and Practice, 4,* 291–306.

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.

Gould, R. A., Otto, M. W., Pollack, M. H., & Liang, Y. (1997). Cognitive-behavioral and pharmacological treatment of generalized anxiety disorder: A preliminary meta-analysis. *Behavior Therapy, 28,* 285–305.

Hansen, N. B., & Lambert, M. J. (in press). An evaluation of the dose-response relationship in naturalistic treatment settings using survival analysis. *Mental Health Services Research*.

Hansen, N. B., Umphress, V., & Lambert, M. J. (1998). The reliability and validity of a short form of the Inventory of Interpersonal Problems. *Journal of Psychoeducational Assessment, 16,* 201–214.

Hewlett, P. S., & Plackett, R. L. (1979). *The interpretation of quantal responses in biology.* London: Edward Arnold.

Hollon, S. D., DeRubeis, R. J., Evans, M. D., Wiemer, M. J., Garvey, M. J., Grove, W. M., & Tauson, V. B. (1992). Cognitive therapy and pharmacotherapy for depression: Singly and in combination. *Archives of General Psychiatry, 49,* 774–781.

Howard, K. I., Kopta, S. M., Krause, M. S., & Orlinsky, D. E. (1986). The dose-effect relationship in psychotherapy. *American Psychologist, 41,* 159–164.

Howard, K. I., Moras, K., Brill, P. L., Martinovich, Z., & Lutz, W. (1996). Evaluation of psychotherapy: Efficacy, effectiveness, and patient progress. *American Psychologist, 51,* 1059–1064.

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

Jacobson, N. S., Follette, W. C., & Revenstorf, D. (1984). Psychotherapy outcome research: Methods for reporting variability and evaluating clinical significance. *Behavior Therapy, 15,* 336–352.

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.

Johnson, R. W. (1965). Number of interviews, diagnosis, and success of counseling. *Journal of Counseling Psychology, 12,* 248–251.

Kadera, S. C., Lambert, M. J., & Andrews, A. A. (1996). How much therapy is really enough? A session-by-session analysis of the psychotherapy dose-effect relationship. *Journal of Psychotherapy Research and Practice, 5,* 132–151.

Klein, D. F., & Smith, L. B. (1999). Organizational requirements for effective clinical effectiveness studies. *Prevention and Treatment, 2,* Article 0002a. Retrieved January 31, 2001, from www.journals.apa.org/prevention/volume2/pre0020002a.html.

Klerman, G. L. (1983). The efficacy of psychotherapy as the basis for public policy. *American Psychologist, 38,* 929–934.

Klosko, J. S., Barlow, D. H., Tassinari, R., & Cerny, J. A. (1990). A comparison of alprazolam and behavior therapy in the treatment of panic disorder. *Journal of Consulting and Clinical Psychology, 58,* 77–84.

Kopta, S. M., Howard, K. I., Lowry, J. L., & Beutler, L. E. (1994). Patterns of symptomatic recovery in psychotherapy. *Journal of Consulting and Clinical Psychology, 62,* 1009–1016.

Kopta, S. M., Lueger, R. J., Saunders, S. M., & Howard, K. I. (1999). Individual psychotherapy outcome and process research: Challenges leading to greater turmoil or a positive transition? In J. T. Spence, J. M. Darlrey, & D. J. Foss (Eds.), *Annual review of psychology* (vol. 50, pp. 441–470). Palo Alto, CA: Annual Reviews.

Ladouceur, R., Dugas, M. J., Freeston, M. H., Leger, E., Gagnon, F., & Thibodeau, N. (2000). Efficacy of a cognitive-behavioral treatment for Generalized Anxiety Disorder: Evaluation in a controlled clinical trial. *Journal of Consulting and Clinical Psychology, 68,* 957–964.

Lambert, M. J., & Bergin, A. E. (1994). The effectiveness of psychotherapy. In A. E. Bergin & S. L. Garfield (Eds.), *The handbook of psychotherapy and behavior change* (4th ed.). New York: Wiley.

Lambert, M. J., Burlingame, G. M., Umphress, V., Hansen, N. B., Vermeersch, D. A., Clouse, G. C., & Yanchar, S. C. (1996b). The reliability and validity of the Outcome Questionnaire. *Clinical Psychology and Psychotherapy, 3,* 249–258.

Lambert, M. J., Hansen, N. B., Umphress, V., Lunnen, K., Okiishi, J., Burlingame, G. M., Huefner, J. C., & Reisinger, C. W. (1996a). *Administration and scoring manual for the Outcome Questionnaire (OQ 45.2).* Wilmington, DE: American Professional Credentialing Services.

Lambert, M. J., Whipple, J. L., Smart, D. W., Vermeersch, D. A., Nielsen, S. L., & Hawkins, E. J. (2001). The effects of providing therapists with feedback on patient progress during psychotherapy: Are outcomes enhanced? *Psychotherapy Research, 11,* 49–68.

Lunnen, K. M., & Ogles, B. A. (1998). A multiperspective, multivariable evaluation of reliable change. *Journal of Consulting and Clinical Psychology, 66,* 400–410.

Maling, M. S., Gurtman, M. B., & Howard, K. I. (1995). The

response of interpersonal problems to varying doses of psychotherapy. *Psychotherapy Research, 5,* 63–75.

McNeilly, C. L., & Howard, K. I. (1991). The effects of psychotherapy: A reevaluation based on dosage. *Psychotherapy Research, 1,* 74–78.

Murphy, G. E., Simons, A. D., Wetzel, R. D., & Lustman, P. (1984). Cognitive therapy and pharmacotherapy: Singly and together in the treatment of depression. *Archives of General Psychiatry, 41,* 33–41.

Nathan, P. E., Stuart, S. P., & Dolan, S. L. (2000). Research on psychotherapy efficacy and effectiveness: Between Scylla and Charybdis? *Psychological Bulletin, 126,* 964–981.

Newman, F. L., & Howard, K. I. (1986). Therapeutic effort, treatment outcome, and national health policy. *American Psychologist, 41,* 181–187.

Nezu, A. M. (1986). Efficacy of a social problem-solving therapy approach for unipolar depression. *Journal of Consulting and Clinical Psychology, 54,* 196–202.

Nezu, A. M., & Perri, M. G. (1989). Social problem-solving therapy for unipolar depression: An initial dismantling investigation. *Journal of Consulting and Clinical Psychology, 57,* 408–413.

Nietzel, M. T., Russell, R. L., Hemmings, K. A., & Gretter, M. L. (1987). Clinical significance of psychotherapy for unipolar depression: A meta-analytic approach to social comparison. *Journal of Consulting and Clinical Psychology, 55,* 156–161.

Norquist, G., Lebowitz, B., & Hyman, S. (1999). Expanding the frontier of treatment research. *Prevention and Treatment, 2,* Article 0001a. Retrieved January 31, 2001, from http://journals.apa.org/prevention/volume2/pre0020001a.html.

Ogles, B. M., Lambert, M. J., & Sawyer, J. D. (1995). Clinical significance of the National Institute of Mental Health Treatment of Depression Collaborative Research Program data. *Journal of Consulting and Clinical Psychology, 63,* 321–326.

Orlinksy, D. E., Grawe, K., & Parks, B. K. (1994). Process and outcome in psychotherapy: Noch einmal. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4th ed., pp. 270–376). New York: Wiley.

Orlinsky, D. E., & Howard, K. I. (1978). The relation of process to outcome in psychotherapy. In S. L. Garfield & A. E. Bergin (Eds.), *Handbook of psychotherapy and behavior change* (3rd ed.). New York: Wiley.

Roth, A., & Fonagy, P. (1996). *What works for whom? A critical review of psychotherapy.* New York: Guilford Press.

Seeman, J. (1954). Counselor judgements of therapeutic process and outcome. In C. R. Rogers & R. F. Dymond (Eds.), *Psychotherapy and personality change.* Chicago: University of Chicago Press.

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.

Shaddish, W. R., Matt, G. E., Navarro, A. M., & Phillips, G. (2000). The effects of psychological therapies in clinically representative conditions: A meta-analysis. *Psychological Bulletin, 126,* 512–529.

Shaddish, W. R., Matt, G. E., Navarro, A. M., Siegle, G., Crits-Christoph, P., Hazelrigg, M. D., Jorm, A. F., Lyons, L. C., Nietzel, M. T., Prout, H. T., Robinson, L., Smith, M. L., Svartberg, M., & Weiss, B. (1997). Evidence that therapy works in clinically representative conditions. *Journal of Consulting and Clinical Psychology, 65,* 355–365.

Shear, M. K., Pilkonis, 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.

Singer, J. D., & Willet, J. B. (1991). Modeling the days of our lives: Using survival analysis when designing and analyzing longitudinal studies of duration and the timing of events. *Psychological Bulletin, 110,* 268–290.

Stevens, S. E., Hynan, M. T., & Allen, M. (2000). A meta-analysis of common factor and specific treatment effects across the outcome domains of the phase model of psychotherapy. *Clinical Psychology: Science and Practice, 7,* 273–290.

Speer, D. C. (1994). Can treatment research inform decision makers? Nonexperimental method issues and examples among older outpatients. *Journal of Consulting and Clinical Psychology, 62,* 560–568.

Standal, S. W., & van der Veen, F. (1957). Length of therapy in relation to counselor estimates of personal integration and other case variables. *Journal of Consulting Psychology, 21,* 1–9.

Steenbarger, B. N. (1994). Duration and outcome in psychotherapy: An integrative review. *Professional Psychology: Research and Practice, 25,* 111–119.

Strassberg, D. S., Anchor, K. N., Cunningham, J., & Elkins, D. (1977). Successful outcome and number of sessions: When do counselors think enough is enough? *Journal of Counseling Psychology, 24,* 477–480.

Strupp, H. H. (1986). Psychotherapy: Research, practice, and public policy (How to avoid dead ends). *American Psychologist, 41,* 120–130.

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

Tang, T. Z., & DeRubeis, R. J. (1999). Sudden gains and critical sessions in cognitive-behavioral therapy for depression. *Journal of Consulting and Clinical Psychology, 67,* 894–904.

Taylor, S. (1996). A meta-analysis of cognitive-behavioral treatments for social phobia. *Journal of Behavioral Therapy and Experimental Psychiatry, 27,* 1–9.

Thompson, L. W., Gallagher, D., & Breckenridge, J. S. (1987). Comparative effectiveness of psychotherapies for depressed elders. *Journal of Consulting & Clinical Psychology, 55,* 385–390.

Tingey, R. C., Lambert, M. J., Burlingame, G. M., & Hansen, N. B. (1996). Assessing clinical significance: Proposed extensions to method. *Psychotherapy Research, 6,* 109–123.

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.

van Balkom, A. J. L. M., Bakker, A., Spinhoven, P., Blaauw, B. M. J. W., Smeenk, S., & Ruesink, B. (1997). A meta-analysis of the treatment of panic disorder with or without agoraphobia: A comparison of psychopharmacological, cognitive-behavioral, and combination treatments. *Journal of Nervous and Mental Disease, 185,* 510–516.

van Oppen, P., de Hann, E., van Balkom, A. J. L. M., Spinhoven, P., Hoogduin, K., & van Dyck, R. (1995). Cognitive therapy and exposure in vivo in the treatment of obsessive-compulsive disorder. *Behaviour Research & Therapy, 33,* 379–390.

Vermeersch, D. A., Lambert, M. J., & Burlingame, G. M. (2000). Outcome Questionnaire: Item sensitivity to change. *Journal of Personality Assessment, 74,* 242–261.

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.

Wampold, B. E. (1997). Methodological problems in identifying efficacious psychotherapies. *Psychotherapy Research, 7,* 21–43.

Wampold, B. E., Mondin, G. W., Moody, M., Stich, F., Benson, K., & Ahn, H. (1997). A meta-analysis of outcome studies comparing bonafide psychotherapies: Empirically, "All must have prizes." *Psychological Bulletin, 122,* 203–215.

Weitz, L. J., Abramowitz, S. I., Steger, J. A., Calabria, F. M., Conable, M., & Yarus, G. (1975). Number of sessions and client-judged outcome: The more the better? *Psychotherapy: Theory, Research, and Practice, 12,* 337–340.

Wiggins, J. G. (1992, May). Practice guidelines: Fundamentally wrong. *APA Monitor,* 3.

Williams, S. L., & Falbo, J. (1996). Cognitive and performance-based treatments for panic attacks in people with varying degrees of agoraphobic disability. *Behaviour Research and Therapy, 34,* 253–264.

Received February 14, 2001; revised May 22, 2001; accepted June 15, 2001