Outcome in Psychotherapy: The Past and Important Advances

Michael J. Lambert
Brigham Young University

Fifty years after the 1963 debate between Strupp and Eysenck, as recorded in their articles in Psychotherapy, it is clear that Eysenck overstated the case against psychoanalysis and dynamic psychotherapy (Bergin, 1971), while inflating the magnitude of improvement in untreated individuals (Lambert, 1976). Eysenck was probably correct about the beneficial effects of behavior therapies, but did not foresee that behavior therapy would be supplanted by cognitive behavior therapies (CBT) and eclectic mixtures of CBT that incorporate elements of eastern religion, humanistic interventions, and psychodynamic constructs. Fortunately, most of the treatments that have been tested in rigorous investigations have been found to be effective, but few have distinguished themselves as uniquely superior. Many of the problems of how to measure the effects of treatment have been solved and suggest that about two thirds of treated individuals improve or recover. This leaves a sizable portion of nonresponding individuals, but emerging methods involving in tracking treatment response are being used to decrease deterioration and enhance positive outcomes.

Keywords: psychotherapy outcome, progress feedback, treatment effects, OQ-45

In the first issue of Psychotherapy, Hans Strupp (1963) wrote a stimulating article that reflected issues of that time period, an auspicious beginning for this new journal. His article was essentially a defense of the effects of psychoanalysis and related analytic/dynamic, as well as eclectic treatments directed at the wholly negative evaluation of them by Hans Eysenck (1952). Strupp wrote persuasively about the ambitious goals of psychoanalysis, the existing evaluations of it, and the degree to which psychoanalysts intended for the effects of their treatment to end with patients being fully functioning individuals whose presenting symptoms disappeared, who were self-understanding and accepting (of both weaknesses and strengths), who were relatively free of ‘talent-crippling inhibitions,’ and who could form and maintain satisfying and lasting interpersonal relationships. The degree of improvement in evaluations of patients was assessed by psychoanalytic therapists who saw patients four times a week over years and who, it can be said, had rather extensive knowledge of their patients and a theoretical and common sense view of ideal mental health. Strupp emphasized the validity of psychoanalytic therapists’ views of their patients’ outcome and strongly advocated for the effectiveness of psychoanalysis based on the scant literature reporting outcomes.

The next year Eysenck (1964) in a later issue of Psychotherapy responded to Strupp’s advocacy and his own appraisal of the effects of psychoanalysis and other psychotherapies reasserting that there was no sound scientific evidence supporting psychoanalysis while advocating both tendencies for patients to heal themselves and the success of behavior therapies. At that point in history, existing studies had not randomly assigned patients to treatment versus no-treatment control groups, and Eysenck capitalized on this fact to assert that patients could be expected to improve over the same period of time without formal treatment—that mental disorders were self-limiting and likely to subside over time. In Strupp’s (1964) rejoinder, he suggested that ample evidence documented the positive effects of psychotherapy and that it was time to move on to more sophisticated questions (such as comparative treatment effects), while also reiterating the importance of broad views of patient benefit: “To be sure, it is important to free a patient from—say—a troublesome phobia, and everyone would agree that symptom relief is a sine qua non in psychotherapy. But, we must not lose sight of the patient’s intrapsychic state—his sense of identity, feeling of worthwhileness as a person, and happiness” (p. 101).

In the context of the initial issue of Psychotherapy and the debate about the effects of psychoanalysis and psychodynamic psychotherapies was the fundamental challenge that behavior therapy posed to psychodynamic thinking and procedures. The growth of learning-based approaches that appeared as early as the 1920s (Jones, 1924; Mowrer & Mowrer, 1938) had not had a dramatic impact on psychotherapy until the publication of Wolpe’s Psychotherapy by Reciprocal Inhibition in 1958. But as Eysenck pointed out, behavior therapy appeared rather hopeful. Still the emergence of cognitive therapy was a natural outgrowth of the limitations of the learning-based approaches, with their emphasis on behavior at the expense of thought, but it also represented dissatisfaction with the effects of psychodynamic treatments. Cognitive therapy was most notably advocated by Ellis (1962) and Beck (1970) and came to the forefront of theory-driven treatments by the mid-1970s with the publication of Beck’s (1976) Cognitive Therapy and the Emotional Disorders. These and related developments, such as the emergence of social learning theory (Bandura, 1969), provided rich contrasts between cognitive theories and treatment methods and carried with them a strong research emphasis.

Of course the debates over the accuracy of Eysenck’s estimates of both the outcomes of the verbal psychotherapies and those of
individuals who did not receive any formal treatment went on for at least another decade, with Bergin’s (1971) comprehensive estimates of change following psychoanalytic psychotherapy (around 85% for those that complete treatment) and Lambert’s (1976) reanalysis of spontaneous remission rates (around 40% improvement). These narrative research reviews and many others suggested that the outcomes for treated individuals were substantial in relationship to the passage of time. Perhaps one of the more convincing studies on the evidence presented by Eysenck was published by McNeilly and Howard (1991) who used Eysenck’s own data to demonstrate a 50% improvement rate in eight sessions of psychotherapy as opposed to a 2% spontaneous remission rate over the same time period, and that patients get 2 years of symptom remission in only 15 sessions of psychotherapy than would happen without psychotherapy.

The Effects of Psychotherapy: Current Findings

In the ensuing years, evidence has mushroomed on a wide variety of older and newer forms of treatment. A search of ISI Web of Knowledge (an Internet journal search engine) reveals that around 60,000 academic articles have been published on psychotherapy research in just the last 30 years. Much of the evidence for the effects of psychotherapy has been summarized in various editions of the Handbook of Psychotherapy and Behavior Change, now out in its 6th edition (Lambert, 2013a). Chapters in the Handbook are organized around systems and modalities of psychotherapy as well as special topics such as the therapist’s contributions to outcomes and those of the client. We now have a mass of well-designed scientific studies that neither Strupp nor Eysenck had access to in 1963 and 1964 when they published their articles in Psychotherapy.

Reanalyses of older reviews, as well as newer meta-analytic reviews of psychotherapy outcome, produce the broad finding of therapy benefit across a range of treatments for a variety of disorders. Indeed, psychotherapy is more effective than many “evidence-based” medical practices, some of which are very costly and produce significant side effects, including almost all interventions in cardiology (e.g., beta-blockers, angioplasty, statins), gynecologic medicine (e.g., calcium and alendronate sodium for osteoporosis), and asthma (e.g., budesonide), influenza vaccine, and cataract surgery, among other treatments (Wampold, 2007). Considering the high burden of illness manifest in psychological disorders, and the fact that the psychotherapies studied last only weeks, the consequences of entering treatment versus having no formal treatment are dramatic. The effect size between treated and untreated individuals produced by quantitative reviews hovers around $d = .75$, leading to an estimate of a general success rate in treated persons of 67% compared with that of 33% for untreated persons over the same period of time. A paradox with these estimates is that the outcome of treated cases weather treated by behavior therapy, cognitive behavior therapy, or other types of psychotherapy (e.g., psychodynamic, emotion-focused psychotherapy) is nearly identical to Eysenck’s estimate of the spontaneous remission rate and therefore Eysenck’s advocacy for behavior therapy is just as vulnerable to his criticisms.

Numerous meta-analytic reviews now consider outcomes of patients with specific disorders. For example, in the past three decades alone, more than 40 meta-analyses (not just studies) have been conducted on the outcomes of patients who have depression (Cuijpers & Dekker, 2005). Results indicate that most psychological treatments that have been studied produce substantial effects, in terms of symptom reduction in depression and increased well-being (Cuijpers, van Straten, van Oppen, & Andersson, 2008), with the number of types of effective psychotherapies rising over time. APA’s Division 12 Task force on empirically supported psychotherapies now lists 12 separate treatments for depression, six with strong evidence and six with lesser evidence (available through American Psychological Association, Division 12). Patients suffering from mood disorders who enter a variety of treatments can expect considerable relief, with the number who will experience a full remission varying with the type of mood disorder and its chronicity. The range of remission probably hovers somewhere between 35% and 70%.

It is fair to say that there is now an abundance of research on treatment outcomes in clinical trials, suggesting that these effects can be achieved in 12–14 sessions of care (hardly an endorsement of psychoanalysis). These individuals will make larger gains than similar individuals on wait-lists, or who receive “placebo” treatments, and they will maintain their gains at follow-ups 2 to 3 years after treatment. It will not generally matter which kind of psychotherapy is offered as long as it is a bona fide theory-driven intervention (Lambert, 2013b).

Many elements of care are shared by diverse treatment orientations and modalities (such as a confidential relationship characterized by high levels of understanding and respect as advocated by client-centered theory, as well as exposure to anxiety-provoking situations), and these common factors seem to loom large in facilitating improved functioning (at least in depression), contributing much more to ultimate outcomes than the kind of specific theory-based interventions that are offered (Cuijpers et al., 2012). Under the somewhat ideal circumstance of clinical trials, it appears that the outcomes are better than those attained in routine care where positive outcomes are found in closer to one third of the patients, and individuals participate for far fewer sessions (around four; Hansen, Lambert, & Forman, 2002). In routine care, it appears that about 50% of treated individuals, including those with mixed and multiple diagnostic features might recover if they received about 18 to 21 sessions of care. About 50% of clients will show reliable improvement following seven sessions of psychotherapy (e.g., Anderson & Lambert, 2001). Although psychoanalytic treatment outcomes continue to be understudied, patients who undergo psychodynamic treatments of much shorter duration (around 20 sessions) appear to fare as well as in other treatments (Barber, Muran, Keefe, & McCarthy, 2013). So Eysenck (1964) was correct in suggesting that psychoanalysis was not efficient, at least for symptom-focused outcomes. The evidence that long-term psychodynamically oriented treatments lead to more substantial changes in personality than other therapies is in doubt given patients’ ability to maintain their treatment gains following much shorter treatments.

Fifty years after the 1963 debate between Strupp and Eysenck, as recorded in their articles in Psychotherapy, it is clear that Eysenck overstated the case against psychoanalysis and dynamic psychotherapy (Bergin, 1971), while inflating the magnitude of improvement in untreated individuals (Lambert, 1976). Eysenck was probably correct about the beneficial effects of behavior therapies, even though they often have very narrow goals—that is,
behavior in specific contexts, and have been supplanted by cognitive behavior therapies and other innovations. There continues to be a debate about the degree to which cognitive therapy can achieve its aims independent of the behavioral component of treatment (Emmelkamp, 2013) and the degree to which behavioral interventions are necessarily independent of the cognitive aspects of treatment, at least for depression (Hollon & Beck, 2013).

Psychotherapy can be quite efficient for a large minority of patients; however, the number of sessions needed for a larger portion (75%) of persons to recover hovers around 50 sessions. In contrast, every theory of behavior change has yet to deal with the surprising and perplexing phenomena of sudden gains (or early dramatic treatment response) in psychotherapy. When patients’ mental health functioning is monitored on a session-by-session basis, it appears that a substantial number of patients (17%–40%) respond to treatments (offered in clinical trials and routine care) much sooner and more substantially than theory would predict. These early responders make up a substantial percentage of recovered/improved individuals at termination and at follow-up years later (e.g., Haas, Hill, Lambert, & Morrell, 2002). Depending on just when such changes occur (with a median session of 5), it appears that they precede many of the interventions thought necessary to bring about change. Such findings certainly present a challenge to psychoanalytic methods and cast serious doubt on the necessity and desirability of prolonged treatment for many individuals.

This phenomenon also presents a challenge to other therapies that advocate a shorter manualized approach to recovery. Tang and DeRubes (1999) have gone to some lengths to attribute sudden gains to components of cognitive behavior therapy with depression, but such an explanation does not take into account the fact that early dramatic improvement occurs in other disorders and without the use of cognitive behavior treatment. The fact that these reported large and lasting benefits occur so early in a wide variety of treatments does suggest that the mechanisms of recovery often involve some kind of dramatic self-reorganization, which are highly dependent on client characteristics. Sudden gains may be the result of a corrective emotional experience, a phrase that came out of psychoanalytic theory that recognized that change was not necessarily incremental (step-by-step), but rather could be sudden, dramatic and surprising in terms of lasting impact.

However, psychoanalysis did not consider dramatic change in early treatment sessions to be credible change, considering it a “flight into health,” a way of escaping the demands of analysis. Instead it appears to be a legitimate common phenomenon across treatment modalities. Clinicians and researchers have noted related phenomena in many forms of psychotherapy (see Hill & Castonguay, 2010; Transformation in Psychotherapy), but little is really known about the causes of corrective experiences and their relationship to rapid dramatic treatment response. An important aspect of the early response is many individuals do not respond in an immediate and dramatic way: their change, if it occurs, takes the form of slower improvement over sessions of treatment.

An often ignored, but critical, consideration in psychotherapy is the degree to which the therapies have negative rather than positive consequences for clients. An estimated 5% to 10% of adult clients participating in clinical trials leave treatment worse off than they began treatment (Lambert & Ogles, 2004). In routine care, the situation is more problematic. Outcomes for >6,000 patients treated in routine practice settings suggest that the clients did not fare as well as those in clinical trials, with deterioration rates as high as 14% in some settings (Hansen et al., 2002). The situation for child psychotherapy in routine care is even more sobering. The small body of outcome studies in community-based usual care settings has yielded a mean effect size near zero (e.g., Weisz, 2004), yet millions of youth are served each year in these systems of care. In a comparison of children being treated in community mental health (N = 936) or through managed care (N = 3075), estimates of deterioration were 24% and 14%, respectively (Warren, Nelson, Mondragon, Baldwin, & Burlingame, 2010).

There is no doubt that all of the deterioration that occurs during the time a patient is in treatment cannot be causally linked to therapist activities. Certainly, a portion of patients are on a negative trajectory at the time they enter treatment and the deteriorating course cannot be stopped. Another subset of patients experience untoward life events that cannot be prevented and that have nothing to do with treatment. A portion of patients are prevented from taking their own lives as a result of effective practices, even if they do not show overall progress. Just as positive psychotherapy outcomes depend largely on patient characteristics, so do the negative changes that occur in patients who are undergoing psychological treatments. Unfortunately, there are reasons to believe that therapists do not recognize client worsening, although there are methods available to help with this problem—a central focus of this article.

Positive as well as negative patient change can be affected by therapist actions and inactions. Research reviews find that the major contribution of the therapist to negative change is usually found in the nature of the therapeutic relationship, with rejections of either a subtle or manifest nature being the root cause (e.g., Lambert, Bergin, & Collins, 1977; Safran, Muran, Samstang, & Winston, 2005).

Unfortunately, it appears that clinicians have an overly optimistic view of their own patients’ progress (Walfish, McAlister, O’Donnell, & Lambert, 2012), estimating very high rates of improvement, and outcomes far superior to their peers. Clinicians overlook negative changes and have a limited capacity to make accurate predictions of the final benefit clients will receive during treatment, particularly with clients who are failing to improve. One study, for example, found that even when therapists were provided with the base rate of deterioration in the clinic where they worked (8%), and were asked to rate each client that they saw at the end of each session (with regards to the likelihood of treatment failure and if the patient was worse off at the current session in relation to their intake level of functioning), they rated only 3 of 550 clients as likely to have a negative outcome and seriously underestimated that their client was worse off at a current session in relation to their intake level of functioning (Hannan et al., 2005). In a separate study, Hatfield, McCullough, Plucinski, and Krieger (2010) found in a retrospective review of case notes of clients who had deteriorated during treatment infrequent mention of worsening even when its degree was dramatic.

Such results are not surprising, given psychotherapist optimism, the complexity of persons, and a treatment context that calls for considerable commitment and determination on the part of the therapist, who actually has very little control over the patient’s life circumstances, decisions, and personal characteristics. Patients’ response to treatment is, especially in the case of a worsening state,
a likely place where outside feedback might have the greatest chance of impact. Helping the therapist become aware of negative change and discussing such progress in the therapeutic encounter are much more likely when formal feedback is provided to therapists. Such feedback helps the client communicate and helps the therapist to become aware of the possible need to adjust treatment, alter or address problematic aspects of the treatment as appropriate (e.g., problems in the therapeutic relationship or in the implementation of the goals of the treatment).

**Measuring, Monitoring, Predicting Treatment Failure, and Using Feedback: Advances in Practice**

A logical procedure for diminishing deterioration and enhancing positive outcomes involves routinely measuring, regularly monitoring, and tracking client treatment response with standardized scales throughout the course of treatment while providing clinicians (and clients) with this information.

**Definitions and Feedback Measures**

Clients can complete a brief measure of their psychological functioning by using standardized rating scales and then this information can be delivered to psychotherapists in real time. Such measures can be regarded as a mental health vital signs lab test. This lab test data can be used to indicate the client’s current level of disturbance in relation to functional and dysfunctional populations, deviations from expected treatment response over the course of psychotherapy, and the consequences of treatment. Collecting this information from the client on a session-by-session basis provides the clinician with a systematic way of monitoring life functioning from the client’s point of view. A brief formal assessment can provide a summary of life functioning that is not otherwise available to the therapist, unless the therapist spends time within the treatment hour to systematically inquire about all the areas of functioning covered by the self-report scale, an activity that detracts from service delivery. Lab test data and information about deviations from an expected treatment response provide novel information to therapists.

This is a quality assurance practice that might be considered a form of *managing* outcomes (Evans, Mellor-Clark, Barkham, & Mothersole, 2006). For many decades, and even to the present day, psychotherapy outcome research, with the notable exception of the behavior therapies, has relied heavily on study designs that measure client outcome at pretreatment and posttreatment. Although such designs have proven beneficial in establishing the general efficacy and effectiveness of the treatments under investigation, they are limited in that outcome data from these studies (because it is collected after termination from treatment) cannot be used to positively influence the treatment process of the individual clients under investigation. Pre- and posttreatment assessments, then, are essentially a “postmortem” analysis of outcome, as clients have already terminated treatment and nothing can be done to improve their outcomes, even if they experienced no change or even deteriorated in treatment. Outcome management extends the practice of measuring and monitoring client progress throughout the course of treatment by then using data collected in real time to positively influence the treatment process and outcome of the clients under investigation. The major advantage of psychotherapy outcome management is that outcome data can be regularly gathered and fed back to clinicians in real time for the purpose of making needed alterations in intervention strategy if clients are either unresponsive to, or deteriorating in treatment.

Several psychotherapy outcome management systems have been developed and implemented in clinical service delivery settings worldwide, such as in the United States (Lambert, Hansen, & Finch, 2001; Lueger et al., 2001; Miller, Duncan, Sorrell, & Brown, 2005), Germany (Kordy, Hannover, & Richard, 2001), the Netherlands (deJong et al., 2007), Australia (Newham, Hook, & Page, 2013), Norway (Anker, Duncan, & Sparks, 2009), and Great Britain (Barkham et al., 2001). Although the specific procedures used in each of these quality management systems vary, a common feature across all of them is the monitoring of client outcome throughout the course of treatment and the use of these data to improve individual client outcomes.

In the remainder of this article, one specific psychotherapy quality management system that has been developed, implemented, and empirically evaluated in multiple investigations is described. The methodology used endeavors to improve psychotherapy outcome by monitoring client progress in relation to expected progress and providing this information to clinicians in order to guide ongoing treatment, particularly for the client who is not having a favorable response to treatment (signal-alarm or Not-On-Track cases). This methodology is an extension of quality assurance and represents one effort to bridge the gap between efficacy and effectiveness research and clinical practice, while enhancing patient outcomes. It is also well suited to models of care in which clinicians attempt to step-up or step-down treatments after assessing patient treatment response (e.g., Otto, Pollack, & Maki, 2000).

**The OQ Psychotherapy Outcome Management System**

Patient outcome can be conceptualized and measured in myriad ways. Strupp and Hadley (1977) proposed a tripartite model for mental health outcomes. They suggested that the three interested parties concerned with therapy outcomes were society, the client, and the mental health professional. Based on this viewpoint, Lambert (1983) suggested the most important aspect of outcome is the subjective experience of the client, including symptoms of anxiety and depression. In addition, the functioning of persons in their social roles (e.g., work) and intimate relationships can be seen as very important for the well-being of the person, family, and society in general. This definition of psychotherapy outcome approximates that used in the early investigations of psychoanalysis where the interest in change goes beyond symptomatic improvements.

As Strupp (1963) noted, psychotherapy outcome is clearly broader than symptom change and includes changes in personality/self-organization. Certainly deeper changes are highly desirable and important to clinicians as well as researchers because they have important implications for lifelong adjustment and coping. The extent to which the OQ-45 and similar measures go beyond merely measuring changes in symptomatic states is open to question. Measures such as the OQ-45, Beck Depression Inventory, Symptom Check List-90, and the like are highly correlated. In my own opinion, changes on these “symptom” scales reflect and are proxies for and estimates of deeper changes in individuals that reflect modifications in foundational views of self and others. But
such a supposition has not been studied and the extent to which it is true has not been estimated and needs clarification in future research.

Quality management systems, as applied in routine practice, demand efficient outcome assessment rather than the more ideal alternative of comprehensive assessment. Outcome measurements that are typically used in efficacy studies often require hours of assessment from multiple perspectives of change. In contrast, outcome-focused research uses weekly assessments with a single brief self-report measure. Thus, assessment in this type of research is much more frequent, with a greater diversity of patients and large final sample sizes, but less comprehensive and lacking multiple perspectives of change. To a large degree, this methodology is essential to use for two reasons. The first is that unlike clinical trials, where the number of treatment sessions a patient is expected to attend is predetermined (usually around 14 sessions), in routine practice, lengths are indeterminate and largely based on client and therapist preferences. In clinical trials, progress can be assessed at preset times during the course of treatment (e.g., at weeks zero, 6, and 12) and especially at termination. In routine care, the outcome measure is collected before each treatment session, so that the effects of the final session are infrequently collected, but outcome data are available for every session up to that point in time. The second reason is related to repeated measurement. Although patients have no problem tolerating a weekly assessment that takes about 5 min, they cannot be expected to undergo lengthier, multiscale, multisource assessments on a frequent basis.

Given the demand for regular and efficient outcome assessment in psychotherapy outcome management, the Outcome Questionnaire-45 (OQ-45; Lambert, Morton, et al., 2004) was developed. It is a 45-item self-report measure designed to assess four domains of client functioning: symptoms of psychological disturbance, particularly anxiety and depression, interpersonal problems, social role functioning, and quality of life (well-being). Consistent with this conceptualization of outcome, the OQ-45 provides a Total Score, based on all 45 items, as well as Symptom Distress, Interpersonal Relations, and Social Role subscale scores. Higher scores on the OQ-45 are indicative of greater levels of psychological disturbance. Research has indicated that the OQ-45 is a psychometrically sound instrument that has been shown to be sensitive to changes in multiple client populations over short periods of time while remaining relatively stable in untreated individuals (Vermeersch, Lambert, & Burlingame, 2000; Vermeersch et al., 2004). It is well suited for assessing initial levels of client distress and tracking client status during and following treatment.

**Defining a Positive and Negative Outcome**

A key element in psychotherapy quality management research is defining and operationalizing the concepts of positive and negative outcome for the individual patient. Jacobson and Truax (1991) offered a methodology by which a client’s change on an outcome measure can be classified in the following categories: recovered, reliably improved, no change, deteriorated. There are two pieces of information necessary in order to make these clients outcome classifications: a Reliable Change Index (RCI) and a normal functioning cutoff score. Clinical and normative data were analyzed by Lambert, Morton, et al. (2004) to establish an RCI and a cutoff score for the OQ-45. Using this information, clients can be placed in the following categories based on the change observed in their OQ scores:

- **Recovered** (i.e., clinically significant change)—Clients whose score decreases by 14 or more points and passes below the cutoff score of 64.
- **Improved** (i.e., reliably changed)—Clients whose score decreases by 14 or more points but does not pass below the cutoff score of 64.
- **No Change**—Clients whose score changes by less than 14 points in either direction.
- **Deteriorated**—Clients whose score increases by 14 or more points.

Support for the validity of the OQ-45’s reliable change and normative cutoff score has been reported by Lutten and Ogles (1998) and Beckstead et al. (2003). Having a method to classify each client’s treatment response is an essential component of outcome management, given that the primary purpose of psychotherapy outcome management is to understand and improve the gains each individual is making during the course of treatment. Furthermore, the ability to classify individual client change further bridges the gap between traditional efficacy and effectiveness studies (that focus on changes made by groups of clients) and clinical practice (Kendall, Marrs-Garcia, Nath, & Sheldrick, 1999).

**Prediction of Treatment Failure**

A core element of outcome management systems is the prediction of treatment failure. To improve outcomes of clients who are responding poorly to treatment, such clients must be identified before termination from treatment, and ideally, as early as possible in the course of treatment. Though many studies have investigated the value of several client, therapist, client–therapist interaction, and extratherapeutic variables in predicting outcome, very few of the variables explored are consistently highly predictive of outcome. Research using the OQ-45 has indicated that the best predictors of outcome are initial severity of distress (i.e., pretreatment OQ-45 total score) and change of score following separate sessions over the course of treatment. In fact, Brown and Lambert (1998) found that pretreatment OQ-45 total score and change scores from Sessions 1 to 3 accounted for approximately 40% of the variance in final outcome, and that after taking these variables into account, all other variables combined (e.g., diagnosis, client demographics, therapist demographics, therapist theoretical orientation, etc.) accounted for less than 1% of the variance in final outcome. In other words, in prior studies using the OQ-45, the best way to predict outcome was to know how distressed clients were prior to treatment and whether or not the changes they made at the session of interest in the treatment process were positive or negative and to what degree.

Given research on the variables most predictive of outcome, an empirically derived signal-alarm system (which plots a statistically generated expected recovery curve for differing levels of pretreatment distress on the OQ-45 and uses this as a basis for identifying clients who are not making expected treatment gains and are at risk of having a poor outcome) was developed to alert clinicians to potential treatment failures. The accuracy of this signal-alarm system has been evaluated in a number of empirical investigations (Ellsworth, Lambert, & Johnson, 2006; Lambert, Whipple, Bishop, et al., 2002; Lutz et al., 2006; Spielmans, Masters, & Lambert, 2002; Spielmans, Masters, & Lambert, 2006).
termination or referral, the current level of misidentification would result in intrusive and even health threatening interventions, the signal-alarm in psychotherapy merely alerts the therapist to the need for reconsidering the value of ongoing treatment, rather than mandating specific changes. Thus, we see the signal-alarm as supporting clinical decision-making, rather than supplanting it. Because the signal-alarm alerts therapists to the possible need for action, rather than triggering a negative chain of events such as termination or referral, the current level of misidentification would seem to be tolerable.

The Provision of Feedback and Clinical Support Tools

The signal-alarm system has been used as an intervention for preventing deterioration and enhancing positive outcomes in clients, in that it alerts clinicians to potential treatment failures and allows them to modify their treatment approach (if they deem that to be appropriate) in an attempt to improve the outcomes of clients who are having a poor response to treatment. Once a client takes the OQ-45, commences treatment, and completes a session of treatment, the signal-alarm system can be used to generate feedback regarding the client’s progress. The feedback to therapists consists of several components, among which are a progress graph that includes all the client’s OQ-45 total scores from pretreatment to the current session and a color-coded message (white, green, yellow, red) that indicates the status of client progress. The specific language of the feedback messages varies not only as a function of client progress, but also as a function of the session at which the feedback is provided (i.e., a red message at Session 2 is not as urgent as a red message at Session 20). An illustrative summary of each feedback message follows:

White Message—“The Client is functioning in the normal range. Consider Termination.”

Green Message—“The rate of change the client is making is in the adequate range. No change in the treatment plan is recommended.”

Yellow Message—“The rate of change the client is making is less than adequate. Recommendations: consider altering the treatment plan by intensifying treatment, shifting intervention strategies, and monitoring progress especially carefully. This client may end up with no significant benefit from therapy.”

Red Message—“The client is not making the expected level of progress. Changes are he or she may drop out of treatment prematurely or have a negative treatment outcome. Steps should be taken to carefully review this case and decide upon a new course of action such as referral for medication or intensification of treatment. The treatment plan should be reconsidered. Consideration should also be given to presenting this client at case conference. The client’s readiness for change may need to be reassessed.”

Over the last 25 years, methodologies have been used in medical research and practice to manage clinical interventions in areas such as drug dosage, diagnosis, and preventive care. These interventions are often used in a stepwise approach that assists physicians in clinical decision-making and provides recommendations to improve the quality of patient health care (Hannan et al., 1998). Similarly, a set of Clinical Support Tools (CST) was developed and integrated into the existing psychotherapy quality management system in an attempt to augment the feedback provided to therapists and further improve outcomes of nonresponding and deteriorating patients (Whipple et al., 2003). As such, the CSTs are intended to be used by therapists only when one of their clients is predicted to have a poor outcome (i.e., when a therapist receives a red or yellow warning message, indicating that client is not responding or deteriorating in treatment).

The CSTs are composed of a problem-solving decision tree designed to systematically direct therapists’ attention to certain factors that have been shown to be consistently related to client outcome in the empirical literature, such as the therapeutic alliance, social support, readiness to change, diagnostic formulation, and need for psychiatric referral. A single measure has been developed (Assessment for Signal Cases; ASC) aimed at assisting therapists to assess the quality of the therapeutic alliance, client readiness for change, client perception of social support, and life events. The CST Manual (Lambert et al., 2007) provides specific intervention strategies that could be used by therapists if problems were detected in the aforementioned domains. These intervention strategies are also included in the OQA Analyst Software (www.OQMeasures.com). When the signal-alarm rules identify a client as Not-On-Track (red or yellow warning), therapists have the option of using the ASC (40 self-report items) and the decision tree that organizes problem-solving hierarchically. Specific items that are problematic are highlighted in the report, and suggested interventions are provided. For example, if a client is identified as at risk for treatment failure and the alliance total score, subscale scores (bond, task, and goal agreement), and specific alliance items are all problematic, the therapist can then examine the list of suggestions for strengthening the relationship (e.g., discuss the clients ratings of the relationship, explore relationship ruptures). If the client’s alliance rating was not below average, the therapist would proceed to evaluation of the client stage of motivation and so forth.

In addition to providing feedback regarding client progress and CSTs to therapists, feedback can also be provided directly to clients. Client feedback messages (i.e., white, green, yellow/red) that correspond to the aforementioned therapist feedback messages have been developed in an effort to directly inform clients of their progress in treatment and enhance client/therapist collaboration in treatment (Harmon et al., 2007; Hawkins, Lambert, Vermeersch, Slade, & Tuttle, 2004). A summary of a client feedback message follows:

Yellow/Red Message—Please note that the following information is based on your responses to the questionnaire that you have completed before each therapy session. It appears that you have not experienced a reduced level of distress. Because you may not be experiencing the expected rate of progress, it is possible that you have even considered terminating treatment, believing that therapy may not be helpful for you. Although you have yet to experience much relief from therapy, it is still early in treatment and there is the potential for future improvement. However, we
urge you to openly discuss any concerns that you may be having about therapy with your therapist because there are strategies that can be used to help you receive the most out of your therapy. It may also require your willingness to complete additional questionnaires that may shed light about why you are not experiencing the expected rate of progress.

Consistent with the findings of previous research (Flowers, 1979; Kivlghan, 1985; Kluger & DeNisi, 1996), the messages designed for patients were a blend of positive and negative language. Effort was made to avoid message content potentially perceived as threatening or discouraging to patients’ self-esteem. Patients were informed of their self-reported level of distress according to the OQ-45, progress since beginning therapy, and likelihood of benefitting from treatment given the present course of progress. Additionally, patients identified as potential treatment failures are encouraged to discuss personal concerns about their progress, alternative courses of action, and goals of therapy with their therapists to further facilitate the collaborative alliance.

The administration of the OQ-45 (whether via paper–pencil or computerized), scoring, application of the signal-alarm system, and generation of feedback reports (for therapists and/or clients) are almost instantaneously processed through software called OQ-Analyst (administration of the measure and generation of the measure takes approximately 5–7 min and is completed before a session usually on a handheld device or online). Accessing the feedback report on the therapist’s personal computer takes approximately 18 s.

**Impact of Feedback on Client Outcome**

Nine controlled studies have been published that examine the effects of providing client progress feedback to therapists and/or clients using the methodology described above (Cris~t-Christoph et al., 2012; Harmon et al., 2007; Hawkins et al., 2004; Lambert, Whipple, et al., 2001; Lambert, Whipple, Vermeersch, et al., 2002; Simon, Lambert, Harris, Busath, & Vazquez, in press; Simon et al., unpublished; Slade, Lambert, Harmon, Smart, & Bailey, 2008; Whipple et al., 2003). Shimokawa, Lambert, and Smart (2010) have reviewed feedback research with patients who are predicted to be treatment failures (as well as patients who appear to be on-track for a positive outcome). This meta/megaanalysis combined data from six well-designed clinical trials that compared treatment-as-usual to feedback-assisted treatments in which the same therapists offered both conditions to over 4,000 patients.

Results indicated that feedback to therapists and patients had a powerful effect over treatment-as-usual with cases that were predicted to be treatment failures (20%–30% of clients). In these off-track clients, the base rate for deterioration at the end of treatment was 20%. Progress feedback with alarm signals indicating a less than expected treatment response reduced deterioration by 50% (to 9%) and increased positive outcomes from 22% in treatment-as-usual to 38% in the feedback condition. In three of the six studies, feedback to therapists included the use of the CST that helped therapists identify reasons for poor therapeutic progress and provided suggestions. This feedback enabled therapists to identify and intervene differently than in treatment-as-usual. The results of this feedback were rather substantial in that it further reduced deterioration rates to 5.5%, with over half of the off-track cases improving or returning to a normal state of functioning at the end of treatment.

Several studies have been conducted since completion of the meta-analytic review suggesting that progress feedback with alerts and problem-solving tools are effective across diverse treatment settings (inpatient, outpatient) and patient samples ranging from the inpatient treatment of eating disorders to substance abuse. The results of these later studies all produced statistically significant improvements in off-track cases compared with treatment-as-usual offered by the same therapists. It is worth pointing out that therapists in the studies we have conducted were not familiar with the feedback procedures before the studies and in most cases seemed skeptical about the value and need for such procedures. In general, therapists were volunteers who agreed to try the tools and were surprised to find that their clients had better outcomes when they attended to the feedback, alarm-signals, and CSTs.

Based on research findings, feedback of the kind just described is considered an evidence-based practice that can be recommended for routine use (see OQ-Analyst at: http://www.nrepp.samhsa.gov/ViewAll). Evidence on alternative feedback systems has been slower to emerge, but also provides positive evidence that progress feedback enhances patient outcomes (Castonguay, Barkham, Lutz, & McAleavey, 2013). Much of the research shows that the effects of feedback are strongest with clients who are struggling to use the therapy to their benefit rather than with those who make consistent progress (about 70%–80% of cases). Unfortunately all patients need to be monitored in order to identify those who significantly deviate from a course of improvement.

As with many innovations in clinical care, clinicians appear hesitant to adopt feedback of the kind just described in routine practice. Besides the natural tendency for providers to stick with the treatments they learned during their training (which does not yet include such feedback), their tendency to overestimate their success rates and underestimate negative outcomes, and their general skepticism of empirical research slows down routine use. Monitoring patient treatment response may be threatening to personal perceptions of exceptional effectiveness held by most clinicians. Collected data on patient outcomes illuminate the differential effectiveness of individual providers. Because such data are normally distributed, and most therapists’ patients have average outcomes, changes in functioning based on standardized scales are usually disappointing to therapists. Documenting differences between providers’ patients’ outcomes allows comparisons to be made between providers, something that clinicians are not used to reconciling with their personal assessment of their effectiveness. It is certainly more comfortable for therapists to practice psychotherapy without formal assessment.

There are a number of limitations in feedback research. The research has relied on single self-report measures to characterize mental health functioning, whereas most psychotherapy outcome research assesses outcome with multiple measures reflecting multiple points of view and multiple types of outcome. More ambitious assessment of outcome could result in more conservative results than we have found through the use of a single self-report measure. This methodological limitation is difficult to overcome because most studies are conducted in routine care clinics with no set treatment lengths or knowledge beforehand on when the final session (and therefore the final assessment) will take place. Another limitation of the meta-analytic results presented here is that...
the studies that were examined were generated by a single research group. Because different measures and feedback methods exist and researchers are beginning to report results, future research may confirm, disconfirm, or partially confirm research by Lambert and his colleagues. At this time, enough studies have been published to suggest that progress feedback is broadly effective as a method of improving patient treatment response with patients who are struggling to improve.

Future Research

Given the preceding review, it follows that I believe a future emphasis is psychotherapy research that concentrates on improving outcomes in real time. Rather than continuing with the kind of research advocated by Strupp and Eysenck in the initial issues of *Psychotherapy* in 1963–64, it is time to focus our attention on improving outcome as psychotherapy is ongoing. In contrast to emphasizing the right psychotherapy for the right disorder, such research assumes that patients do need an empirically supported psychotherapy that is not working for them. Instead, it is important to identify negatively responding and nonresponding individuals as early as possible in treatment, quickly analyze reasons for these failures, and alternatives or modifications to the clinical approach. Our initial efforts have been to measure generic factors (the therapeutic alliance, motivational problems, and outside factors such as social supports), but this work is just at its inception. There may be some theory-specific, clinic-specific, and population-specific that would also be highly effective and extend this work. For example, in our local clinic group, therapies abound and are easily accessible for patients with social support problems. So if problems with social support are identified, a referral can be made. Our measure, however, does not identify whether the problem in this area is due to needed more contact with others or whether a group that focusses on social skill deficiencies would be more helpful. In either case, identifying what is going wrong and creating within a specific clinic some intervention possibilities that routinely work is a creative clinical and research task. A recent examination of treatment failures in cognitive behavior therapies provides an example of a similar approach. Dimidjian and Hollon (2011) brought together a series of articles on treatment failure, with many of the contributors analyzing the reasons for failure and steps that you could be taken with similar cases in the future. There is much to be learned from studying treatment failure.

Anyone familiar with the scientific study of psychotherapy treatment effects is aware of the importance of measurement—the operational definition of impact and change. Just how many patients are estimated to have improved following a course of treatment is directly related to what is measured and with what instruments. This is an area where psychotherapy research has made clear advances. In contrast to the research results discussed by Strupp and Eysenck in 1963/4 that often relied on therapist ratings, contemporary research has used multiple standardized measures from a variety of sources to provide comprehensive assessments of change. Ogles (2013), based on a review of outcome studies published over a 3-year period, found that the typical study used multiple measures, with an average of 3.89 measures of outcome per study, with a range from 1 measure to 14. Twenty-four percent of the outcome studies included an outcome from a significant other such as spouse, parent, and so forth. Importantly, only 5% of the studies included an unstandardized measure (one with no prior reliability or validity data). A total of 435 unique outcome measures were used in the 163 studies he examined, indicating both the variable ways in which the effects of psychotherapy are assessed and the lack of agreement about the best measures to capture change. In stark contrast to early practices involving evaluation of psychoanalytic psychotherapy where therapists were central in evaluations, therapist assessments of change are now rare. This is consistent with recommendations from outcome researchers.

In the area of assessment, considerable work needs to be done to help us understand clinically significant change based on popular measures of outcome. For example, just how do change scores on self-report measures of symptoms reflect broader changes in the “deeper” aspects of client functioning? Certainly we have come a long way since the first issue of *Psychotherapy* appeared and we have further to go, but we can be proud that our treatments rest on a firm empirical foundation and that scientific efforts in combination with clinician contributions continue to improve clinical services and modify our theoretical understanding of psychotherapy and the change process.

References


Received September 24, 2012

Accepted September 25, 2012