This article was downloaded by: [Curtin University Library]
On: 29 July 2013, At: 02:16
Publisher: Routledge
Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK

Psychotherapy Research
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tpsr20

Implementing routine outcome monitoring in clinical practice: Benefits, challenges, and solutions

James F. Boswell [a], David R. Kraus [b], Scott D. Miller [c] & Michael J. Lambert [d]

[a] Department of Psychology, University at Albany, State University of New York, Albany, NY, USA

[b] Outcome Referrals, Framingham, MA, USA

[c] ISTC, Chicago, IL, USA

[d] Department of Psychology, Brigham Young University, Provo, UT, USA

Published online: 26 Jul 2013.

To cite this article: Psychotherapy Research (2013): Implementing routine outcome monitoring in clinical practice: Benefits, challenges, and solutions, Psychotherapy Research, DOI: 10.1080/10503307.2013.817696

To link to this article: http://dx.doi.org/10.1080/10503307.2013.817696

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the “Content”) contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at http://www.tandfonline.com/page/terms-and-conditions

Psychotherapy Research, 2013
http://dx.doi.org/10.1080/10503307.2013.817696

Implementing routine outcome monitoring in clinical practice: Benefits, challenges, and solutions

JAMES F. BOSWELL[1], DAVID R. KRAUS[2], SCOTT D. MILLER[3], & MICHAEL J. LAMBERT[4]

[1]*Department of Psychology, University at Albany, State University of New York, Albany, NY, USA;* [2]*Outcome Referrals, Framingham, MA, USA;* [3]*ISTC, Chicago, IL, USA &* [4]*Department of Psychology, Brigham Young University, Provo, UT, USA*

(Received 3 April 2013; revised 12 June 2013; accepted 18 June 2013)

Downloaded by [Curtin University Library] at 02:16 29 July 2013

Abstract
This article reviews the benefits, obstacles, and challenges that can hinder (and have hindered) implementation of routine outcome monitoring in clinical practice. Recommendations for future routine outcome assessment efforts are also provided. Spanning three generations, as well as multiple developed tools and approaches, the four authors of this article have spent much of their careers working to address these issues and attempt to consolidate this learning and experience briefly here. Potential "elephants in the room" are brought into the discussion wherever relevant, rather than leaving them to obstruct silently the field's efforts. Some of these topics have been largely ignored, yet must be addressed if we are to fulfill our promise of integrating science and practice. This article is an attempt to identify these important issues and start an honest and open dialogue.

Keywords: routine outcome monitoring; feedback; science-practice integration; dissemination and implementation

> "Ideas are a dime a dozen. People who implement them are priceless."
>
> Mary Kay Ash

In 1662, the Flemish medicinal chemist Jan Baptist van Helmont proposed what is considered the first randomized clinical trial (RCT) in history. He challenged conventional wisdom held as fact by practitioners of his day. Venesection, better known as bloodletting, was a major therapeutic tool. The procedure, believed effective for the remedy of a variety of pathologic conditions, was bolstered by an elaborate rationale, system of differential diagnosis, and specific tools and techniques. Confident that science would prevail over tradition, he offered a wager of 300 Florins—approximately 60,000 modern US dollars—to any of his contemporaries willing to test their methods against his:

> Let us take out of the Hospitals, out of the Camps, or from elsewhere, 200, or 500 poor People, that have Fevers, Pleurisies, etc. Let us divide them in Halfes, let us cast lots, that one half of them may fall to my share and the other to yours; I will cure them without bloodletting . . . but you do as ye know . . . we shall see how many funerals both of us shall have." (Harford, 2011, p. 121).

No record exists of anyone accepting van Helmont's bet. What is known is that he was arrested, interrogated, condemned, and confined by authorities of the day. Bloodletting continued for another 200 years. Even if his contemporaries had agreed to put their practices to the test and been proven wrong, it would be a mistake to assume rapid acceptance and adoption of the data. Then as now, implementation significantly lags behind discovery (McHugh & Barlow, 2012). Despite a culture favorable to the scientific method, a literate population, and the ability to transmit information at speeds unimaginable in van Helmont's age, available evidence indicates that as much as two decades can pass before discoveries are integrated into clinical practice (Brownson, Kreuter, Arrington, & True, 2006).

The purpose of this paper is, in part, to identify the barriers to implementation—important factors that when ignored silently impede progress. In particular, we focus on the implementation of routine outcome monitoring (ROM). Following a brief introduction and making a case for the impact

---

Correspondence concerning this article should be addressed to James F. Boswell, University at Albany, State University of New York, Department of Psychology, Albany, NY, USA. Email: jboswell@albany.edu

© 2013 Society for Psychotherapy Research

of ROM, we identify philosophical and practical barriers to such monitoring. Finally, we offer some potential solutions to impediments encountered by front-line clinicians and healthcare systems when putting ROM into practice.

Outcome Monitoring

In 1996, Howard, Moras, Brill, Martinovich, and Lutz first suggested using session-to-session measures of client progress to evaluate and improve treatment outcome. Their approach differed from traditional efficacy and effectiveness research, which focuses on the average response of participants in either experimental or naturalistic settings. As a complement to traditional nomothetic approaches, these researchers proposed directing increased attention to a more idiographic approach, asking, "Is this treatment, however constructed, delivered by this particular provider, helpful to this client at this point in time?"

Since Howard et al.'s (1996) pioneering work, over a dozen RCTs and several meta-analyses have been published (e.g., Shimokawa, Lambert, & Smart, 2010). These studies provide strong empirical support for ROM. When implemented, the risk of patient deterioration is significantly decreased. At the same time, effect sizes are enhanced, and in some extreme cases tripled (Anker, Duncan, & Sparks, 2009; Kraus, Castonguay, Boswell, Nordberg, & Hayes, 2011). Moreover, technological advances now enable practitioners to quickly and efficiently administer reliable and valid measures, track progress, and receive individualized feedback for their clients in real time (Barkham, Mellor-Clark, Connell, & Cahill, 2006; Kraus, Wolf, & Castonguay, 2006; Lambert, 2012; Miller, Duncan, Sorrell, & Brown, 2005). As can be seen in the companion papers of this series, ROM has been the focus of practice-oriented research in diverse naturalistic settings (Castonguay, Pincus, & McAleavey, in press; Fernández-Álvarez, Gómez & García, in press; Homqvist, Philips, & Barkham, in press; McAleavey, Lockard, Castonguay, Hayes, & Locke, in press; Strauss et al., in press; West et al. in press).

If adoption and implementation were a simple matter of combining evidence with a practicable methodology, then ROM would not only be known and accepted, but widely used. This is not consistently the case (Miller, Hubble, Chow, & Seidel, 2013). Surveys spanning different countries indicate that few clinicians actually employ ROM in their day-to-day work (Gilbody, House, & Sheldon, 2002; Hatfield & Ogles, 2004; Zimmerman & McGlinchey, 2008). Furthermore, although the collection of routine data has received more attention from service organizations and provider groups in recent years, there is considerable variability in the implementation, sustainability, and subsequent use of routine outcome data. Despite indications of low utilization, Bickman (2000) found that a large percentage of therapists held interest in receiving regular reports of client progress. Subsequently, Hatfield and Ogles (2004) conducted a survey with a national sample of licensed psychologists to investigate this discontinuity. As before, clinicians expressed interest in having reliable outcome information. Among the reasons given by those choosing not to use outcome measures, the top were "practical (e.g., cost and time) and philosophical (e.g., relevance) barriers" (p. 485).

Downloaded by [Curtin University Library] at 02:16 29 July 2013

Outcome Monitoring Systems

Drapeau (2012) identified 10 measures/systems for tracking mental health changes in routine care and had the authors of each system describe their purposes and implementation procedures. Obviously, there are many unique ways to monitor patients' mental health functioning over the course of treatment. In order to maximize the discussion of lessons learned about implementing ROM in clinical care, discussion will be limited to the following systems, several of which some of the present authors have developed: PCOMS ICCE, TOP, CORE, and OQ Systems.

The *Partners for Change Outcome Management System: International Center for Clinical Excellence* (PCOMS ICCE) includes measures of progress and the therapeutic alliance. Scales are available for adults, youth, and children and have been translated into 25 different languages. The brevity of these measures makes them ideal for use at every session. Following an independent review, PCOMS ICCE was recently listed on the National Registry of Evidence-based Programs and Practices. Based on three RCTs included in a recent meta-analysis, Lambert and Shimokawa (2011) found an effect size of $g = .53$ for the treatment-as-usual vs. feedback groups, with an odds ratio indicating that the feedback groups were 3.5-times more likely to achieve *reliable change*.

Another ROM and feedback measurement system is the *Treatment Outcome Package* (TOP; Kraus, Seligman, & Jordan, 2005), with different versions for adults, adolescents, and children. The TOP system was built on the recommendations of the Core Battery Conference (Strupp, Horowitz, & Lambert, 1997), and is unique in its multidimensional focus (Kraus, Boswell, Wright, Castonguay, & Pincus, 2010). The TOP factor structure allows clinicians to track change across 12 different dimensions of behavior, symptoms, quality of life and functioning. The adult version of the instrument assesses dimensions that include substance abuse, depression, panic, psychosis, mania,

Downloaded by [Curtin University Library] at 02:16 29 July 2013

suicide, violence, sleep, quality of life, and social, work and sexual functioning. Recent research using the TOP has focused on identifying clinicians' relative strengths and weaknesses in specific problem domains and with particular types of clients (Kraus et al., 2011), including the examination of outcomes for clients with complex comorbidity (Nordberg et al., 2010).

A third system, used extensively in the United Kingdom, is the *Clinical Outcomes in Routine Evaluation* (CORE) system developed by Barkham and colleagues (2001, 2006). Similar to the TOP, the CORE has been used not only to track client progress but to benchmark patient outcomes for use at clinic and system levels. Thus, administrators in conjunction with clinicians can identify underperforming units and track improvements following modification to routine care. At relatively little cost, therapists can formally measure, monitor, and track patient self-reported well-being, and predict as well as improve final treatment response, especially with clients who worsen during the course of treatment (for a detailed description of the CORE, see Homqvist et al., in press).

The final group of measures considered is the *Outcome Questionnaire System* (OQ System). Much of the pioneering research on the subject of ROM was done using the scales developed by Lambert and colleagues (Lambert et al., 1996). The OQ system offers different outcome measures for adults, children, adolescents, and clients considered seriously mentally ill. Additional features of the OQ and supporting research are discussed in the section that follows.

Each of the above-named systems has similarities and differences, as well as strengths and weaknesses (Youn, Kraus, & Castonguay, 2012), and adequate empirical support. PCOMS is the briefest of the group, thereby potentially facilitating greater acceptance by practitioners. Similarly, the CORE is widely used in Europe and has two abbreviated versions for use in routine clinical practice. The TOP and OQ systems are both relatively longer, and the OQ is the most researched tool of the lot. Involving the greatest number of items and multidimensional feedback, the TOP was recently tested by the Annie E. Casey Foundation with users in Ohio. The range of administration times was 3 to 5 minutes. The TOP has also been a data linchpin for innovative and productive community practice research networks (PRNs; e.g., Castonguay et al., 2010) and data-driven treatment initiatives (see Adelman, 2005). That said, choosing the right tool for a particular setting involves working to strike a balance between the competing demands of validity, reliability, and feasibility as well as meeting individual clinician or agency preferences and needs (Miller, Duncan, & Hubble, 2004). Following this introduction, we now briefly make the case for ROM before discussing philosophical and practical barriers to such monitoring, as well as offering some solutions to impediments encountered by front line clinicians and healthcare systems when putting ROM into practice.

Why Do We Need Outcome Monitoring?

In a recent article, Youn, Kraus, and Castonguay (2012) identified several benefits to ROM. For example, the detection of even slight improvements can reassure skeptical clients that they are making recognizable progress in treatment and further improve the therapeutic alliance. Furthermore, ROM can provide therapists with "off track" alerts indicating that the current course of treatment may be ineffective or even harmful. Even in controlled treatment settings, rates of deterioration and nonresponse are not insignificant. Hansen, Lambert, and Forman (2002) examined a representative sample of randomized clinical trial outcomes based on 89 treatment comparisons (mostly CBT) and reported an average of 57% to 67% recovered or improved after receiving an average of 13 sessions of treatment. These outcomes were contrasted with those found in over 6,000 clients who participated in routine care that lasted an average of four sessions with patients ranging from those treated in community mental health centers to those being seen in outpatient clinics. Rates of improvement/recovery averaged 35% and deterioration varied from a low of 3.2% to a high of 14%, with an average rate of 8%. As Hansen et al. (2002), point out, even when an EST is offered to individuals who have the same disorder and see therapists who have been carefully selected, monitored, and supervised, 30% to 50% of patients fail to respond to treatment.

The situation for child and adolescent outcome in routine care is also sobering. In a comparison of children being treated in community mental health ($N = 936$) or through managed care ($N = 3{,}075$), estimates of deterioration were 24% and 14%, respectively (Warren, Nelson, Mondragon, Baldwin, & Burlingame, 2010). This means that even if there were a right treatment or "best practice" for an individual, we would need to identify patients who are failing to respond to this treatment before they left our care. Furthermore, increased attention to deterioration in treatment may be warranted given the high rates of treatment dropout observed in clinical practice. It is estimated that 40–60% of children and adolescents discontinue treatment prematurely (Kazdin, 1996; Wierzbicki & Pekarik, 1993); many of these dropouts are probably due to perceived lack of benefit from treatment. With regard to measuring treatment response in child and adolescent psychotherapy, Kazdin (2005) noted that "such information would be enormously helpful if used to monitor and evaluate treatment in clinical practice" (p. 555).

Unfortunately, clinicians' view of their own clients' outcome is much more positive. Walfish, McAlister, O'Donnell, and Lambert's (2012) survey of clinicians suggests that clinicians estimate that about 85% of their clients improve or recover. In addition, they have the common impression that they are unusually successful, with 90% rating themselves in the upper quartile and none seeing themselves as below average in relation to their peers. Another serious problem in practice is that doubt exists regarding the ability of clinicians to identify clients during the course of therapy who ultimately deteriorate and to note worsening during treatment as a warning sign of deterioration and treatment failure (Hannan et al. 2005; Hatfield, McCullough, Plucinski, & Krieger, 2010). Clinicians could benefit from using tracking systems because of their likely overly optimistic estimates of their clients' outcome and their inability to predict treatment failure, specifically, reliable negative change.

Predicting Negative Change

One core element of some feedback systems is their proven ability to predict treatment failure, risk of hospitalization, or other negative outcomes. In order to improve outcomes of clients who are responding poorly to treatment, such clients must be identified before termination, and ideally as early as possible in the course of treatment. Systems employ a variety of methods to predict treatment failure, e.g., the OQ system plots a statistically generated expected recovery curve for differing levels of pre-treatment distress 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 (not-on-track cases). The accuracy of this signal-alarm system has been evaluated in a number of empirical investigations (e.g., Ellsworth, Lambert, & Johnson, 2006; Lambert, Whipple, Bishop, et al., 2002) that suggest 85% to 100% of those who eventually deteriorate can be identified before they leave treatment. This rate of recognition far exceeds clinical judgment alone (Hannan et al., 2005).

The Benefits of Routine Outcome Monitoring: Beyond Prediction

In the most recent meta-analytic review of one ROM system, Shimokawa and colleagues (2010) re-analyzed the combined dataset ($N = 6151$) from six

OQ-45 feedback studies published to that date (Harmon et al., 2007; Hawkins, Lambert, Vermeersch, Slade, & Tuttle, 2004; Lambert et al., 2001; Lambert, Whipple, Vermeersch, et al., 2002; Slade, Lambert, Harmon, Smart, & Bailey, 2008; Whipple et al., 2003). Each of the studies evaluated the effects of providing feedback about each client's improvement through the use of progress graphs and warnings about clients who were not demonstrating expected treatment responses. The six studies shared many design and methodological features: (a) consecutive cases seen in routine care regardless of client diagnosis or co-morbid conditions (rather than being disorder specific); (b) random assignment of clients to experimental conditions (various feedback interventions) and treatment as usual (TAU) conditions (no feedback) was made in four of the six studies, while reasonable measures were taken in two studies to ensure equivalence in experimental and control conditions at pre-treatment; (c) psychotherapists provided a variety of theoretically guided treatments, with most adhering to cognitive behavioral and eclectic orientations and fewer representing psychodynamic and experiential orientations; (d) a variety of therapist experience— post-graduate therapists and graduate students each accounted for about 50% of clients seen; (e) *therapists saw both experimental (feedback) and no feedback cases*, thus limiting the likelihood that outcome differences between conditions could be due to therapist effects; (f) the outcome measure as well as the methodology rules/standards for identifying signal-alarm clients (failing cases) remained constant; (g) the length of therapy (dosage) was determined by client and therapist rather than by research design or arbitrary insurance limits.

The meta-analysis (Shimokawa et al., 2010) involved both intent-to-treat (ITT) and efficacy analyses on the effects of various feedback interventions in relation to TAU (treatment without feedback) on clients who were predicted to have a negative outcome. When the not-on-track feedback group was compared to the not-on-track TAU group, the effect size for post-treatment OQ score difference averaged a $g = .53$. These results suggest that the average at risk client whose therapist received feedback was better off than approximately 70% of at risk clients in the no feedback condition. In terms of the clinical significance at termination, 9% of those receiving feedback deteriorated while 38% achieved clinically significant improvement. In contrast, among at risk clients whose therapists did not receive feedback, 20% deteriorated while 22% clinically significantly improved. When the odds of

Downloaded by [Curtin University Library] at 02:16 29 July 2013

Downloaded by [Curtin University Library] at 02:16 29 July 2013

deterioration and clinically significant improvement were compared, results indicated those in the feedback group had less than half the odds of experiencing deterioration while having approximately 2.6-times higher odds of experiencing reliable improvement.

The OQ feedback system went beyond progress feedback by asking clients who were predicted to deteriorate to complete a 40-item measure of the therapeutic alliance, motivation, social supports, and recent life events. Therapists were provided with feedback on these domains, a problem-solving decision tree, and intervention suggestions to assist them in resolving issues that may be causing clients to have a negative treatment response. Together this intervention was referred to as a *Clinical Support Tool.* When the outcome of clients whose therapist received the Clinical Support Tool feedback were compared to the treatment-as-usual clients, the effect size for the difference in mean post-treatment OQ scores was $g = 0.70$. These results indicate that the average clients in the Clinical Support Tool feedback group, who stay in treatment to experience the benefit of this intervention, are better off than 76% of clients in treatment-as-usual. The rates of deterioration and clinically significant improvement among those receiving Clinical Support Tools were 6% and 53%, respectively. The results suggest that clients whose therapists used Clinical Support Tools with off-track cases have less than a fourth the odds of deterioration, while having approximately 3.9-times higher odds of achieving clinically significant improvement.

The applications of not-on-track predictive modeling are potentially far-reaching. For example, TOP outcome data have been tied to prospective health-plan claims data to build and test algorithms that predict near-future psychiatric and substance abuse hospitalizations. These predictive models allow resources to be diverted to prevent costly and potentially life-threatening incidents (McAleavey, Nordberg, Kraus, & Castonguay, 2012). Additional advantages of some outcome monitoring approaches include the general documentation of change for accountability purposes, assistance in formulating treatment plans with built-in evaluations of success, as well as improved communication between therapist and client, especially as it pertains to the discontinuation and prolonging of treatment (Youn et al., 2012).

The preceding information suggests an answer to the question posed: Why would clinicians want to formally monitor client treatment response? Because the research evidence strongly supports the conclusion that it is in clients' (and, thus, therapists') best interests to do so. However, there are many obstacles to the implementation of ROM in routine treatment settings. Examples of such obstacles, and a few strategies that have been used to overcome them, are discussed in the next section.

Obstacles and Challenges to Routine Data Collection

While we view ROM and feedback as having clear benefits for clients, therapists, and healthcare, there are many obstacles and challenges to routine data collection, spanning individual and systems levels. Furthermore, some obstacles and challenges may be unique to the healthcare systems of specific countries. It is clear that researchers would like access to real-world outcome data and that clinicians hold access rights to these data. Cooperation is obviously required, and it is imperative that researchers understand the needs of clinicians, who often need help with answers and solutions to time-critical, sometimes life-threatening, problems. Understanding the clinician's challenges and context is vital. Using some of the primary reasons for non-ROM use reported by Hatfield and Ogles (2004) as broad categories ("practical" and "philosophical"), we now briefly review several obstacles and challenges to data collection that we have encountered.

Practical Obstacles

Financial burden. At the systems level, we believe that there is a financial "elephant in the room." In most healthcare settings, when the industry decides that a type of laboratory data (e.g., an X-ray) is important for delivering quality care, it is billable and reimbursed (either by a third-party payer or by the national health system). Physicians do not pay out of pocket for X-rays. If an orthopedist purchases an X-ray machine for her office, she bills the insurance company each time it is used. This type of reimbursement does not occur with the collection and reporting of outcome data (through a diagnostic or testing code), despite the fact that it clearly improves clients' lives. Importantly, the National Business Group on Health recommended that patient monitoring as a lab test be reimbursed (Finch & Phillips, 2005); however, the insurance industry failed to act on this recommendation. We presume that if the industry had followed this recommendation, many more behavioral health providers would begin routinely collecting data on their clients. If routine monitoring were reimbursed at $25/month, the process would not be a financial burden to providers. We can find no articles that bemoan the fact that physicians have to pay for non-reimbursable tests for their patients. Furthermore, we can think of

Downloaded by [Curtin University Library] at 02:16 29 July 2013

no parallel medical example where physicians are expected to pull out their wallets and pay for a routine procedure that has been demonstrated in clinical trials (as reviewed above) to improve health and outcomes.

Time burden. Time is everything to a busy behavioral health provider, and providers are often overscheduled. Many providers have reported feeling pressured to see clients even though they are probably too tired, distracted, or sick to give their clients their full effort (see Pope, Tabachnick, & Keith-Spiegel, 1987). Similarly, if providers had more time to make phone calls to collaborating providers, as well as request and read previous medical records that relate not only to previous behavioral health treatments but to interrelating co-morbid medical conditions, routine treatment would likely result in better outcomes. In our experience, it often seems that simply having the time to reflect on a case and integrate ideas learned from continuing education or recent journal articles is a luxury that most providers do not have.

Routine data collection must be placed within this context; otherwise clinicians may rightly feel dismissed. Depending on the method or system of measurement, potential time burdens include: administering questionnaires (providing and collecting); scoring and interpretation of results; creating a report and providing feedback to the client; establishing a tracking and reminder system that reliably prompts when follow-up assessments should be conducted (if not collected at every session). Some systems also require time for aggregate reporting, benchmarking, risk adjusting, quality improvement integration, and report interpretation aids. To most agencies, and certainly clinicians in private practice, these challenges can be daunting. The development of software systems can alleviate some of these burdens, yet the time commitment is real.

It is also important to remember that the time burden can affect creativity and self-learning. The average clinician may not have time to research different outcome monitoring systems (e.g., how does one pick the "right" measure) or how to integrate an outcome monitoring and feedback system into his or her practice. Additionally, even if a clinician already is routinely collecting data, he or she may not have much time to reflect on its use or impact. Each of the present authors has had experiences where previously "resistant" clinicians apologized for initially assuming that outcome monitoring would be a distraction rather than an aid to treatment; literally calling to apologize and to relate a story of some unexpected clinical utility. However, these calls are more the exception rather than the rule—likely due to the diminished opportunity to reflect.

Multiple stakeholders with different needs. It is difficult enough to develop an outcome monitoring and feedback system that meets the needs of just one key stakeholder, and we cannot forget that clients will always be the principal stakeholder. Clients want the highest quality of care, but they also want relatively brief scales with a high degree of face validity. Clinicians need actionable data that informs treatment planning and warns of off-track or emergent problems, as well as data that helps to foster and support the development of a stronger therapeutic alliance. Administrators also need actionable data, yet with minimal disruption and cost. In the United States, health plans have a contractual obligation to manage the care secured by the purchasers of healthcare (e.g., employers), yet need to abide by privacy and security laws. A poorly designed system that does not meet the needs of all stakeholders is sure to be short lived (Fixsen, Naoom, Blase, Friedman, & Wallace, 2005).

Turnover. For provider agencies that have already begun to implement ROM, the departure of the right clinical or administrative ("local") champion can cause years of data collection and integration to unravel within weeks. For example, the new CEO might ask, "Why are we spending this money handing out questionnaires? Are we required to do this?" After this, it does not take long for routine data collection and feedback efforts to be jettisoned. In addition, clinical staff turns over in many locations at an increasingly rapid pace, causing the need for ongoing training, education, and seemingly never-ending "buy-in" discussions. As such, turnover not only effects motivation, but also greatly impacts the sustained implementation of routine data collection procedures.

Philosophical Obstacles

Perception that outcome assessment is different from other assessment. As much as we may try, psychologists will likely never be able to create the "perfect" assessment tool: Perfect in the sense that it is perfectly reliable, valid, appropriate and sufficient (i.e., relevant) for each individual client. Because behavioral health providers see a wide range of individuals presenting with diverse and comorbid problems, clinicians are often skeptical about the relevance and utility of outcome measures, despite very solid empirical support for the reliability and validity of many different measures. Despite the accuracy of actuarial prediction, all assessment tools

Downloaded by [Curtin University Library] at 02:16 29 July 2013

must be interpreted within a context and while taking into consideration multiple sources of information (McAleavey et al., 2012). A psychologist would (hopefully) never diagnose a child with attention-deficit hyperactivity disorder (ADHD) on the basis of a single self-report scale. However, this is not because the self-report scale is necessarily faulty or provides no incremental validity. Simply put, outcome assessment is assessment, and should be approached as such. One manifestation of this issue emerges when a client may believe it is in his or her best interest to understate (or overstate) their problems and produce inaccurate ratings on feedback systems. These systems and their usefulness in treatment are predicated on accurate self-reporting of levels of disturbance and corresponding changes. Some clinicians may falsely interpret this as a problem with the measure or feedback system. However, this may actually provide an opportunity for the clinician to take other information into consideration and discuss the discrepancy openly with the client. Such moments can provide glimpses into the unique world-view of the patient and the window through which the therapist and patient can look together (Youn et al., 2012).

Additionally, Youn et al. (2012) note that therapists may resist integrating routine outcome assessments due to concerns that clients will refuse to cooperate and that the process of outcome assessment will interfere with forming a therapeutic alliance. We are aware of no studies (single-case or otherwise) that support this proposition. On the contrary, Youn et al. note that well-introduced outcome procedures may actually improve the therapeutic alliance and clients often welcome the opportunity to track their progress in treatment.

Fear and mistrust. Most formal monitoring of patient mental health is being imposed on clinicians by systems of care that have come to realize that such practices can enhance patient outcomes or because they believe that the data can be used to increase the quality of care offered to patients through accountability measures. This development provides clinicians with further reasons to resist monitoring—we do not like external control and management. Clinicians don't like "big brother" and perhaps with good reason. What will happen with the data? Who will have access to the data? Will it be used to cut reimbursement? Will the insurance companies use it to deny other care? These concerns extend to clinicians and employers. For example, how will the hospital or community mental health center use data as it relates to performance reviews, supervision, and raises? In addition, most clinicians are not used to anything intruding on the therapy hour, including the prying eyes of assessments that will be potentially seen by someone else. Nor do they like having their patients' outcomes compared to those of other therapists (Okiishi et al., 2006). Clinicians may fear that the results of outcome monitoring will reveal them as incompetent (Youn et al., 2012). Obviously, this can raise anxieties and most clinicians will need time to find the process worthy of their trust.

Privacy and ethics. Confidentiality is an absolutely crucial element of psychotherapy, and anything perceived to potentially breach confidentiality can and should be met with skepticism. Ethics charges have been levied against programs that are not perceived to ensure confidentiality (Koocher & Keith-Spiegel, 2008) and the American Psychological Association (APA) has set forth some basic recommendations that health plans are encouraged to follow (APA, 2009). For example, if an outcome system tracks an individual with an identifiable field, such as a name or social security number, simply complying with HIPAA requirements may not be enough unless the outcome/data collection organization has contractual business associate agreements with providers, health plans and any other covered entity involved in the program.

Attempts to Overcome Obstacles

The present authors are keenly aware of the practical and philosophical obstacles described above because we have directly faced and continue to wrestle with them. Through these experiences, we have attempted several strategies to address these difficulties, which have resulted in some success. For example, in order to address potential concerns about the reliability and validity of outcome measures, research is continually being conducted to improve our measurement and feedback systems and provide additional support for psychometric properties and clinical utility. As more data are collected, our predictive analytics will continue to improve.

We have also noted that the perceived time and energy burden is a frequently expressed concern. In order to reduce this time burden and make outcome assessment more user-friendly, each of us has been involved in building software applications that help to ameliorate time burdens. Systems have been built to take over the reminder and administrative burdens by "pushing" questionnaires to consenting patients at required intervals, a process that can eliminate the need for paper or computer solutions in waiting rooms.

The issues of trust and confidentiality are complex and multifaceted. In our experience, clinician trust

Downloaded by [Curtin University Library] at 02:16 29 July 2013

is enhanced when there is a high degree of transparency and everything is spelled out in writing. We have also learned that researchers' attempts to impart the "wisdom of routine outcome monitoring" are far less effective than the wisdom imparted by fellow clinicians who have used the particular outcome monitoring system of interest. It is through direct clinical experience and by sharing these experiences (e.g., through vignettes) that other clinicians begin to seriously entertain the potential benefits. We envision a book or online repository of these clinician-driven case vignettes, which can be mined by other clinicians and administrators who feel as if they are "herding cats."

Furthermore, in our experience, agencies that have successfully solved the turnover challenge (and loss of a "local champion") have done two things. First, outcome data and monitoring become a routine agenda item of every board meeting so that a new executive will learn from their board why the data are invaluable. Second, the champion filling the role of "outcomes project manager" (usually a clinical or quality director) has created a new clinician initial training protocol that introduces staff to the requirements, expectations and benefits of outcome management.

Recommendations for Future Efforts

Despite considerable evidence that psychotherapists are not alert to treatment failure (e.g., Hannan et al., 2005; Hatfield, et al, 2010), and strong evidence that clinical judgments are usually inferior to actuarial methods (Garb, 2005), therapists' confidence in their clinical judgment alone stands as a barrier to implementation of monitoring and feedback systems. We advocate the employment of real-time client feedback procedures to compensate for therapists' limited ability to accurately detect client worsening in psychotherapy. Clients are best served when standardized procedures are used by clinicians to monitor clients' response to psychotherapy and satisfaction with the therapy relationship. Such feedback improves psychotherapy outcomes and certainly does so for clients at risk of deterioration or dropout. Based on our experience developing and implementing routine outcome assessment tools and systems (our successes and failures), we offer the following specific recommendations for future efforts in this crucial area. These recommendations intentionally span practical and conceptual issues related to adoption, implementation, and sustainability, in order to be applicable to a wider range of current and potential ROM adopters.

Consider Additional Incentives

Theories of motivation indicate that goals are more likely to be pursued and achieved when there is frequent, objective feedback that is tied to specific behaviors, and when those specific behaviors (or one's performance) is clearly tied to an outcome that is perceived to be worthwhile (Kluger & DeNisi, 1996; Koestner, 2008). The feedback research reviewed above underscores the value of ROM and clearly demonstrates the direct benefits to both clinicians and their clients. The results are robust and clinically meaningful—at the level of the individual client. However, Hatfield and Ogles (2007) found that clinicians were more likely to report using outcome measures when this practice was linked to payer factors. In the context of healthcare reform, calls have been made to link reimbursement rates with use of evidence-based practices and so-called "proof of" or likelihood of "benefit" (Diamond & Kaul, 2009; Rosenthal, 2008). One method for enhancing the adoption of outcome monitoring and feedback systems may be to incentivize such routine outcome data collection. The obvious incentives would be monetary; however, other potentially useful incentives could be utilized. For example, client referrals could be made based on whether or not a clinician routinely monitors client outcomes or has a track record of helping clients with similar profiles; or clinicians could obtain CEUs through their involvement in data collection and feedback-seeking. Importantly, starting in 2015, the Medicare Physician Quality Reporting System (PQRS) will begin *penalizing* providers with reduced reimbursement rates if data are not reported on designated service measures.

Data Collection Process Must be Simple and Minimally Disruptive

The ROM process must be made as simple as possible and minimally disruptive to clinicians' routine practice. This includes minimal disruption to the clinician *as well as the client* (Slade, Thornicroft, & Glover, 1999). Completion of relevant self-report measures typically consumes no more than 5 minutes of a client's time, which we believe meets the spirit of minimal disruption. Even when the implementation of an outcome monitoring system is relatively straightforward, clinicians' needs must be considered. For example, will implementing this procedure necessitate any formal training, which might also include a tutorial on how to interpret and make use of specific feedback (e.g., Clinical Support Tools)? Clinicians should know how to speak with their clients about the importance of

Downloaded by [Curtin University Library] at 02:16 29 July 2013

outcome monitoring because certain clients feel it may be in their interest to understate (or overstate) their problems and produce inaccurate ratings on feedback systems. Generally speaking, we advocate using electronic versions of feedback systems that expedite and ease practical difficulties. Fortunately, the software for the OQ and TOP, for example, can provide instantaneous feedback to clinicians. If the client completes these measures immediately prior to the scheduled psychotherapy session, electronic feedback is available within seconds to the therapist prior to beginning that session.

Increased Flexibility

Depending on the individual clinician and practice setting, interests, needs, and resources are likely to vary. For example, sophisticated electronic/online outcome assessment, tracking and feedback systems exist; however, many clinicians may be unwilling or able to use such a system (e.g., due to lack of a computer in their office, concerns about confidentiality). In such cases, alternative methods of data collection and scoring may be preferred. For example, a clinician who primarily works with older adults might have concerns about the use of technology, making a paper-pencil option more preferable in some instances. Although different assessment methods may theoretically introduce another source of error variance, studies comparing paper and computer or internet scale administration have found high correspondence between the two assessment methods (Cook et al., 2007; Merten & Ruch, 1996; Ogles, France, Lunnen, Bell, & Goldfarb, 1998; Peterson, Johannsson & Carlsson, 1996). Maintaining flexibility is also consistent with Bohanske and Franczak's (2010) recommendation to think "evolution" when implementing an innovation, rather than "revolution." At the same time, however, failing to take advantage of modern information technology cannot be excused by lack of familiarity or discomfort if progress is to be made.

Transparent and Non-Hierarchical Approach

Outcome monitoring and feedback is an evidence-based practice, and can be treated as such (McHugh & Barlow, 2012). In fact, the effect sizes reported in experiments comparing feedback with treatment as usual exceed those typically reported in comparative outcome studies that justify the use of evidence-based therapies and other evidence-based practices that are widely advocated (Lambert, 2013). With regard to evidence-based practices, skepticism and "top-down" approaches to dissemination and implementation are significant contributors to the science-practice "chasm" (Castonguay, Barkham, Lutz, & McAleavey, 2013). Clinician mistrust can be attenuated by increased transparency regarding the nature, goals, costs, and benefits of monitoring client outcomes. Although we view the costs as very minimal in comparison to the clear benefits outlined above, dissemination and implementation models and research indicate that changing clinician behavior (or anyone's behavior for that matter) is a difficult process, particularly in the beginning (see McHugh & Barlow, 2012; Riemer, Rosof-Williams, & Bickman, 2005; Rogers, 2003), and the ROM implementation process can span several years (Fixsen et al., 2005).

There are many different approaches to changing human systems (Chin & Benne, 1969). The so-called empirical-rational approach assumes that individuals are rational and will follow the data if it makes sense and seems to suit their self-interests; "If we publish enough studies with large effect sizes, then clinicians will start monitoring outcomes." The power-coercive approach assumes that change is facilitated through external leverage where power is the primary factor (e.g., economic pressures, sanctions); "Clinicians will start monitoring outcomes because we will mandate that they do so or refuse to reimburse for services." Alternatively, the normative-educative approach is more a systems approach, where the individual is seen as actively searching to satisfy needs and interests. In contrast with the inherent passivity in a power-coercive approach, the individual takes specific action to advance interests and goals—change is seen as a development of new habits and a realization of values; "If we together discuss why this would be important for clinicians self-interests and the interests of their clients, then we can begin working together to devise a plan." This approach is more participatory, less hierarchical (top-down), and involves clinicians in the planning, which is likely to increase motivation (Koestner, 2008; Latham & Locke, 2006; Locke & Latham, 2006). This also improves transparency.

In line with this, concerns about confidentiality, who will have access to the data collected (during and after a specified collection period) and how, specifically, the data are to be used should be addressed and made clear from the outset. For example, clinicians should be made aware that in addition to tracking (and, therefore, improving) outcomes, they can objectively assess their effectiveness with particular groups of clients once enough data have been collected. Administrators should be made aware that outcomes data can be used as a needs assessment to direct future training initiatives at their mental health center. Additionally, although individual clinician "buy in" is paramount, the context of care delivery is typically much more

Downloaded by [Curtin University Library] at 02:16 29 July 2013

complex. Many initiatives, particularly those at the organizational level (e.g., introducing an electronic records system that integrates ROM), will need to involve administrators, supervisors, non-direct service staff (e.g., receptionists), and in some cases, third-party payers or referral sources.

Identify a Local Champion and Ensure this Role is Always Filled

One of the most consistently reported factors in the successful adoption (Greenhalgh, Robert, Macfarlane, Bate, & Kyriakidou, 2004) and implementation of evidence-based practices, including ROM, is the presence of a "local champion" (Aarons, 2005; Marty, Rapp, McHugo, & Whitley, 2007; Trauer, Gill, Pedwell, & Slattery, 2006). A local champion is a prominent and well-respected individual (or set of individuals) within a setting, organization, or culture who has had positive experiences with a process or intervention and who takes enthusiastic responsibility for assisting in the adoption, implementation, and sustainability of a specific program or intervention. For example, a well-liked and respected veteran clinician or administrator at a clinic who has had positive experiences with routine outcome monitoring and feedback might function as a local champion. This person is not only important for her or his technical and procedural expertise (e.g., can assist in the day-to-day training or troubleshooting when a question or problem arises), but because this individual can engender a more open or positive attitude toward adoption and implementation.

While this problem is likely endemic, anecdotally, each of the present authors has been directly involved in an initiative that eventually lost momentum or failed due to the absence or loss of a local champion. Consequently, we advocate identifying several potential local champions, if possible, rather than a single local champion, at least one of whom holds an administrative position (where applicable). Many clinics and community mental health centers, particularly those specializing in substance abuse, have a high staff turnover rate (Eby, Burk, & Maher, 2010). Once a local champion leaves a setting, particularly if the attitudes of other clinicians are equivocal and the monitoring system is relatively new, sustainability will likely be undermined (Rogers, 2003).

Supplement Outcome Feedback with Clinical Support Tools

Although feedback alarm-signals alone have been shown to be effective in decreasing deterioration and increasing response rates, clinicians are likely to benefit from additional information or decision-support tools in their attempts to help clients who are at risk. Consequently, we recommend that outcome monitoring be supplemented with clinical support tools. As suggested by the general literature on feedback and the evidence presented here, problem-solving and decision-enhancement tools prove helpful to clinicians and, most importantly, clients whose treatment response is in doubt.

Use of Benchmarking and Risk-Adjustment

In order to maximize the validity and "informational payload" of data collected through routine outcome monitoring, benchmarks at the inter- and intraindividual levels should be established and, when possible, data should be risk-adjusted. Statistical benchmarking enhances the utility and interpretability of a measure or indicator (Tremblay, Hevner, & Berndt, 2012), and allows for the identification of types of patients with whom a clinician is particularly effective (or ineffective) as well as those clinicians who generally evidence better outcomes (e.g., outcomes achieved by the top 10% of clinicians; Weissman et al., 1999). There is some evidence that outcomes are enhanced when individuals receive feedback that compares their performance with a statistically derived benchmark, rather than feedback on individual performance (e.g., Kiefe et al., 2001). Risk-adjustment algorithms are important for increasing the comparability of collected data because they adjust for patient characteristics (i.e., case-mix variables, or case-mix adjustment) that could account for differences between clinicians or mental health centers (data collected at centers that serve individuals with differing levels of baseline severity or functional impairment). Of course, the relevance of risk-adjustment likely depends on the aim and scope of the project. For example, if a large mental health provider organization was interested in identifying which of its satellite clinics was achieving particularly good outcomes and which was achieving particularly poor outcomes, risk-adjustment would be important for increasing the fairness and validity of comparisons. Once identified using a valid, data-driven approach, more resources can be allocated to the satellite center that may be struggling, or clinicians from the satellite clinic that is performing particularly well can offer insights into what seems to be working for them.

Conduct More Basic Research on Adoption and Implementation

Basic research on ROM and feedback is essential. Although some attention has been paid to general models of dissemination and implementation (e.g.,

Downloaded by [Curtin University Library] at 02:16 29 July 2013

Fixsen et al., 2005; Rogers, 2003), more needs to be learned about the adoption, implementation, and sustained use of outcome monitoring and feedback systems, as this information may enhance future adoption and compliance, and, therefore, further improve outcomes. For example, more research is needed to empirically test implementation models, such as the one proposed by Fixsen et al. (2005) involving the following stages: *exploration*, *installation*, *initial implementation*, *full implementation*, and *sustainability*. More basic research is also sorely needed on the factors (e.g., participant factors, organizational factors, training factors) that facilitate or inhibit the adoption, implementation, and sustainability of ROM (see De Jong, van Sluis, Nugter, Heiser, & Spinhoven 2012 for an example of such research).

Foster the Scientist-Practitioner Professional Identity

Narrowly defining the scientist-practitioner as a clinician who delivers empirically supported treatments (ESTs) may unintentionally weaken clinicians' identification with this model in a manner that is specifically relevant to valuing outcome monitoring and feedback-seeking. That is, in line with the notion of the local clinical scientist (see Stricker & Trierweiler, 1995), a scientist-practitioner is always an active participant (rather than an observer or mode of delivery) in all relevant domains and sees each clinical encounter as an opportunity to gather "local data" that will not only be relevant with this client, but with future clients. Consequently, training programs should instill the value of collecting routine data, on both process and outcome, and using this information to inform case conceptualization and treatment planning (Castonguay, Boswell, Constantino, Goldfried, & Hill, 2010). In addition, training faculty would do well to encourage an openness to receiving progress feedback (Boswell & Castonguay, 2007), as well as encourage the use of outcomes data to answer clinically relevant research questions early on in training (Castonguay et al., in press).

Conclusion

Movement toward evidence-based and outcome-driven healthcare decision making and resource allocation is an inevitability. Despite the obstacles and challenges outlined above, we believe that most clinicians would agree with the spirit of routine outcome monitoring given that the ultimate goal is reducing suffering and improving individuals' quality of life. How this is precisely achieved and its success will be dependent on the active and open collaboration of researchers and clinicians. Although Jan Baptist van Helmont's challenge was apparently unmet, more attention and resources are being directed toward the rigorous testing of patient-centered approaches to treatment decision making and care, such as the Association for Health Care Research and Quality (AHRQ) and the Patient-Centered Outcome Research Institute (PCORI). We are all excited about the potential improvements in care and the benefits to the individual client and therapist.

References

Aarons, G. A. (2005). Measuring provider attitudes toward evidence-based practice: Consideration of organizational context and individual differences. *Child and Adolescent Psychiatric Clinics of North America*, *14*, 255–271. http://dx.doi.org/10.1016/j.chc.2004.04.008

Adelman, R. (2005). Reducing adolescent clients' anger in a residential substance abuse treatment facility. *Journal on Quality and Patient Safety*, *31*, 325–327.

American Psychological Association. (2009). Criteria for the Evaluation of Quality Improvement Programs and the Use of Quality Improvement Data. *American Psychologist*, *64*, 551–557. doi:10.1037/a0016744

Anker, M. G., Duncan, B. L., & Sparks, J.A. (2009). Using client feedback to improve couple therapy outcomes: a randomized clinical trial in a naturalistic setting. *Journal of Consulting & Clinical Psychology*, *77*, 693–704. doi:10.1037/a0016062

Barkham, M., Margison, F., Leach, C., Lucock, M., Mellor-Clark, J., Evans, C., ... McGrath, G. (2001). Service profiling and outcomes benchmarking using the CORE-OM: Towards practice-based evidence in the psychological therapies. *Journal of Consulting and Clinical Psychology*, *69*, 184–196. doi:10.1037/0022-006X.69.2.184

Barkham, M., Mellor-Clark, J., Connell, J., & Cahill, J. (2006). A core approach to practice-based evidence: A brief history of the origins and applications of the CORE-OM and CORE System. *Counselling & Psychotherapy Research*, *6*, 3–15. doi:10.1080/14733140600581218

Bickman, L. (2000). Summing up program theory. *New Directions for Evaluation*, 103–112. doi:10.1002/ev.1186

Bohanske, R. T., & Franczak, M. (2010). Transforming public behavioral health care: A case example of consumer-directed services, recovery, and the common factors. In B.L. Duncan, S.D. Miller, B.E. Wampold, & M.A. Hubble (Eds.), *The heart and soul of change: Doing what works in therapy* (2nd ed., pp. 299–322). Washington DC: American Psychological Association.

Boswell, J. F., & Castonguay, L. G. (2007). Psychotherapy training: Suggestions for core ingredients and future research. *Psychotherapy: Theory, Research, Practice, Training*, *44*, 378–383. doi:10.1037/0033-3204.44.4.378

Brownson, R. C., Kreuter, M. W., Arrington, B. A., & True, W. R. (2006). Translating scientific discoveries into public health action: How can schools of public health move us forward? *Public Health Reports*, *121*, 97–103.

Castonguay, L. G., Barkham, M., Lutz, W., & McAleavey, A. A. (2013). Practice-oriented research: Approaches and applications. In M.J. Lambert (Ed.), *Bergin and Garfield's handbook of therapeutic change* (6th ed.). New York: Wiley & Sons.

Castonguay, L. G., Boswell, J. F., Constantino, M. J., Goldfried, M.R., & Hill, C.E. (2010). Training implications of harmful

Downloaded by [Curtin University Library] at 02:16 29 July 2013

effects of psychological treatments. *American Psychologist*, *65*, 34–49. doi:10.1037/a0017330

Castonguay, L. G., Boswell, J. F., Zack, S. E., Baker, S., Boutselis, M. A., Chiswick, N. R. . . . Grosse Holtforth, M. (2010). Helpful and hindering events in psychotherapy: A practice research network study. *Psychotherapy: Theory, Research, Practice, Training*, *47*, 327–344. doi:10.1037/a0021164

Castonguay, L. G., Pincus, A. L., & McAleavey, A. A. (in press). Practice-research networks in training clinics. *Psychotherapy Research*.

Chin, R., & Benne, K. D. (1969). General strategies for effecting changes in human systems. In W.G. Bennis, K.D. Benne, & R. Chin (Eds.), *The planning of change* (pp. 32–59). New York: Holt, Rinehart & Winston.

Cook, I. A., Balasubramani, G. K., Eng, H., Friedman, E., Young, E. A., Martin, J., . . . Wisniewski, S. R. (2007). Electronic source materials in clinical research: Acceptability and validity of symptom self-rating in major depressive disorder. *Journal of Psychiatry Research*, *41*, 737–743. doi:10.1016/j.jpsychires.2006.07.015

De Jong, K., van Sluis, P., Nugter, M. A., Heiser, W. J., & Spinhoven, P. (2012). Understanding the differential impact of outcome monitoring: Therapist variables that moderate feedback effects in a randomized clinical trial. *Psychotherapy Research*, *22*, 464–474. doi:10.1080/10503307.2012.673023.

Diamond, G. A., & Kaul, S. (2009). Evidence-based financial incentives for healthcare reform: Putting it together. *Circulation: Cardiovascular quality and outcomes*, *2*, 134–140. doi:10.1161/CIRCOUTCOMES.108.825695.

Drapeau, M. (2012). The value of tracking in psychotherapy. *Integrating Science & Practice*, *2*, 2–6.

Eby, L. T., Burk, H., & Maher, C. P. (2010). How serious of a problem is staff turnover in substance abuse treatment? A longitudinal study of actual turnover. *Journal of Substance Abuse Treatment*, *39*, 264–271. doi:10.1016/j.jsat.2010.06.009.

Ellsworth, J. R., Lambert, M. J., & Johnson, J. (2006). A comparison of the Outcome Questionnaire-45 and Outcome Questionnaire-30 in classification and prediction of treatment outcome. *Clinical Psychology & Psychotherapy*, *13*, 380–391. doi:10.1002/cpp.503

Fernández-Álvarez, H., Gómez, B., & García, F. (in press). Bridging the gap between research and practice in a clinical and training network. *Psychotherapy Research*

Finch, R. A., & Phillips, K. (2005). Center for Prevention and Health Services. *An employer's guide to behavioral health services: A roadmap and recommendations for evaluating, designing, and implementing behavioral health services*. Washington DC: National Business Group on Health.

Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M., & Wallace, F. (2005). Implementation research: A synthesis of the literature. Tampa, FL: University of South Florida, Louis de la Parte Florida Mental Health Institute, The National Implementation Research Network (FMHI Publication #231).

Garb, H. N. (2005). Clinical judgment and decision making. *Annual Review of Clinical Psychology*, *1*, 67–89. doi:10.1146/annurev.clinpsy.1.102803.143810.

Gilbody, S., House, A., & Sheldon, T. (2002). Psychiatrists in the United Kingdom do not use outcomes measures. *British Journal of Psychiatry*, *180*, 101–103. doi:10.1192/bjp.180.2.101

Greenhalgh, T., Robert, G., Macfarlane, F., Bate, P., & Kyriakidou, O. (2004). Diffusion of innovations in service organizations: Systematic review and recommendations. *Milbank Quarterly*, *82*, 581–629. doi:10.1111/j.0887-378X.2004.00325.x.

Hannan, C., Lambert, M. J., Harmon, C., Nielsen, S. L., Smart, D. W., Shimokawa, K., & Sutton, S. W. (2005). A lab test and algorithms for identifying clients at risk for treatment failure. *Journal of Clinical Psychology: In Session*, *61*, 155–163. doi:10.1002/jclp.20108

Hansen, N. B., Lambert, M. J., & Forman, E. V. (2002). The psychotherapy dose-response effect and its implications for treatment delivery services. *Clinical Psychology: Science and Practice*, *9*, 329–343. doi:10.1093/clipsy.9.3.329

Harford, T. (2011). *Adapt: Why success always starts with failure*. London: Macmillan.

Harmon, S. C., Lambert, M. J., Smart, D. W., Hawkins, E. J., Nielsen, S. L., Slade, K., & Lutz, W. (2007). Enhancing outcome for potential treatment failures: Therapist/client feedback and clinical support tools. *Psychotherapy Research*, *17*, 379–392. doi:10.1080/10503300600702331

Hatfield, D., McCullough, L., Plucinski, A., & Krieger, K. (2010). Do we know when our clients get worse? An investigation of therapists' ability to detect negative client change. *Clinical Psychology & Psychotherapy*, *17*, 25–32.

Hatfield, D. R., & Ogles, B. M. (2004). The current climate of outcome measures use in clinical practice. *Professional Psychology: Research and Practice*, *35*, 485–491. doi:10.1037/0735-7028.35.5.485

Hatfield, D. R., & Ogles, B. M. (2007). Why some clinicians use outcome measures and others do not. *Administrative Policy Mental Health and Mental Health Services Research*, *34*, 283–291. doi:10.1007/s10488-006-0110-y

Hawkins, E. J., Lambert, M. J., Vermeersch, D. A., Slade, K., & Tuttle, K. (2004). The effects of providing patient progress information to therapists and patients. *Psychotherapy Research*, *14*, 308–327. doi:10.1093/ptr/kph027

Homqvist, R., Philips, B., & Barkham, M. (in press). Developing and delivering practice-based evidence: Observations, tensions, and challenges. *Psychotherapy Research*.

Howard, K. I., Moras, K., Brill, P. L., Martinovich, Z., & Lutz, W. (1996). Evaluation of psychotherapy. *American Psychologist*, *51*, 1059–1064. doi:10.1037/0003-066X.51.10.1059

Kazdin, A. E. (1996). Dropping out of child therapy. Issues for research and implications for practice. *Clinical Child Psychology and Psychiatry*, *1*, 133–156. doi:10.1177/1359104596011012

Kazdin, A. E. (2005). Evidence-based assessment for children and adolescents: Issues in measurement development and clinical application. *Journal of Clinical Child & Adolescent Psychology*, *34*, 548–558. doi:10.1207/s15374424jccp3403_10

Kiefe, C. I., Allison, J. J., Williams, O. D., Person, S. D., Weaver, M. T., & Weissman, N. W. (2001). Improving quality improvement using achievable benchmarks for physician feedback: a randomized controlled trial. *JAMA*, *285*, 2871–2879. doi:10.1001/jama.285.22.2871

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, *119*, 254–284. doi:10.1037/0033-2909.119.2.254.

Koestner, R. (2008). Reaching one's personal goals: A motivational perspective focused on autonomy. *Canadian Psychology*, *49*, 60–67. doi:10.1037/0708-5591.49.1.60.

Koocher, G. P., & Keith-Spiegel, P. (2008). *Ethics in psychology and the mental health professions: Standards and cases*. New York: Oxford University Press.

Kraus, D. R., Boswell, J. F., Wright, A. G. C., Castonguay, L. G., & Pincus, A. L. (2010). Factor structure of the treatment outcome package for children. *Journal of Clinical Psychology*, *66*, 627–640.

Kraus, D. R., Castonguay, L. G., Boswell, J. F., Nordberg, S. S., & Hayes, J. A. (2011). Therapist effectiveness: Implications for accountability and patient care. *Psychotherapy Research*, *21*, 267–276. doi:10.1080/10503307.2011.563249

Kraus, D. R., Seligman, D., & Jordan, J. R. (2005). Validation of a behavioral health treatment outcome and assessment tool

Downloaded by [Curtin University Library] at 02:16 29 July 2013

designed for naturalistic settings: The Treatment Outcome Package. *Journal of Clinical Psychology*, *61*, 285–314. doi:10.1002/jclp.
20084

Kraus, D. R., Wolf, A., & Castonguay, L. (2006). A kinder philosophy to the management of outcomes. *Psychotherapy Bulletin*, *41*, 23–31.

Lambert, M. J. (2010). *Prevention of treatment failure: The use of measuring, monitoring, and feedback in clinical practice*. Washington DC: American Psychological Association Press.

Lambert, M. J. (2012). Helping clinicians to use and learn from research-based systems: The OQ-Analyst. *Psychotherapy*, *49*, 109–114. doi:10.1037/a0027110.

Lambert, M. J. (2013). The efficacy and effectiveness of psychotherapy. In M.J. Lambert (Ed.) *Bergin & Garfield's Handbook of psychotherapy & behavior change* (6th ed.). New York: Wiley.

Lambert, M. J., Burlingame, G. M., Umphress, V., Hansen, N. B., Vermeersch, D. A., Clouse, G. C., & Yanchar, S. C. (1996). The reliability and validity of the outcome questionnaire. *Clinical Psychology & Psychotherapy*, *3*, 249–258. doi:10.1002/(SICI)1099-0879(199612)3:4<249::AID-CPP106>3.0.CO;2-S

Lambert, M. J. & Shimokawa, K. (2011). Collecting client feedback. In J.C. Norcross (Ed.), *Psychotherapy relationships that work* (2nd ed., pp. 203–223). New York: Oxford University Press.

Lambert, M. J., Whipple, J. L., Bishop, M. J., Vermeersch, D. A., Gray, G. V., & Finch, A. E. (2002). Comparison of empirically derived and rationally derived methods for identifying clients at risk for treatment failure. *Clinical Psychology and Psychotherapy*, *9*, 149–164. doi:10.1002/cpp.333

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 client progress during psychotherapy: Are outcomes enhanced? *Psychotherapy Research*, *11*, 49–68. doi:10.1080/713663852

Lambert, M. J., Whipple, J. L., Vermeersch, D. A., Smart, D. W., Hawkins, E. J., Nielsen, S. L., & Goates, M. K. (2002). Enhancing psychotherapy outcomes via providing feedback on client progress: A replication. *Clinical Psychology and Psychotherapy*, *9*, 91–103. doi:10.1002/cpp.324

Latham, G. P., & Locke, E. A. (2006). Enhancing the benefits and overcoming the pitfalls of goal setting. *Organizational Dynamics*, *35*, 332–340. doi:10.1016/j.orgdyn.2006.08.008

Locke, E. A., & Latham, G. P. (2006). New directions in goal-setting theory. *Current Directions in Psychological Science*, *15*, 265–268. doi:10.1111/j.1467-8721.2006.00449.x.

Marty, D., Rapp, C. A., McHugo, G. J., & Whitley, R. (2007). Factors influencing consumer outcome monitoring in implementation of evidence-based practices: Results from the National EBP Implementation Project. *Administration and Policy in Mental Health*, *35*, 204–211. doi:10.1007/s10488-007-0157-4.

McAleavey, A. A., Lockard, A. J., Castonguay, L. G., Hayes, J. A., & Locke, B.D. (in press). Research by, with, and for clinicians: Lessons from a large-scale PRN for University and College Counseling Centers. *Psychotherapy Research*.

McAleavey, A. A., Nordberg, S. S., Kraus, D. R., & Castonguay, L.G. (2012). Errors in treatment outcome monitoring: Implications of multidimensional and general measurements for real-world psychotherapy. *Canadian Psychology*, *53*, 105–114. doi:10.1037/a0027833

McHugh, R. K., & Barlow, D. H. (2012). *Dissemination and implementation of evidence-based psychological interventions*. New York: Oxford University Press.

Merten, T., & Ruch, W. (1996). A comparison of computerized and conventional administration of the German versions of the Eysenck Personality Questionnaire and the Carroll Rating Scale for Depression. *Personality and Individual Differences*, *20*, 281–291. doi:http://dx.doi.org/10.1016/0191-8869(95)00185-9

Miller, S. D., Duncan, B. L., & Hubble, M. A. (2004). Beyond integration: The triumph of outcome over process in clinical practice. *Psychotherapy in Australia*, *10*, 32–43.

Miller, S. D., Duncan, B. L., Sorrell, R., & Brown, G. S. (2005). The Partners for Change Outcome Management System. *Journal of Clinical Psychology: In-Session*, *61*, 199–208. doi:10.1002/jclp.20111.

Miller, S. D., Hubble, M. A., Chow, D. L., & Seidel, J. A. (2013). The outcome of psychotherapy: Yesterday, today, and tomorrow. *Psychotherapy*, *50*, 88–97. doi: 10.1037/a0031097

Nordberg, S. S., Boswell, J. F., Kraus, D. R., Castonguay, L. G., Hayes, J., & Wampold, B. (2010, June). *Therapist effectiveness treating depression with and without co-morbid substance abuse*. Paper presented at the Society for Psychotherapy Research Conference, Asilomar, CA.

Ogles, B. M., France, C. R., Lunnen, K. M., Bell, M. T., & Goldfarb, M. (1998). Computerized depression screening and awareness. *Community Mental Health Journal*, *34*, 27–38. doi:10.1023/A:1018760128239.

Okiishi, J., Lambert, M. J., Eggett, D., Nielsen, S. L., Dayton, D.D., & Vermeersch, D.A. (2006). An analysis of therapist treatment effects: Toward providing feedback to individual therapists on their clients' psychotherapy outcome; *Journal of Clinical Psychology*, *62*, 1157–1172. doi:10.1002/jclp.20272

Peterson, L., Johannsson, V., & Carlsson, S. (1996). Computerized testing in a hospital setting: Psychometric and psychological effects. *Computers in Human Behavior*, *12*, 339–350. doi:http://dx.doi.org/10.1016/0747-5632(96)00012-X

Pope, K. S., Tabachnick, B. G., & Keith-Spiegel, P. (1987). Ethics of practice: The beliefs and behaviors of psychologists as therapists. *American Psychologist*, *42*, 993–1006. doi:10.1037/0003-066X.42.11.993

Riemer, M., Rosof-Williams, J., & Bickman, L. (2005). Theories related to changing clinician practice. *Child and Adolescent Psychiatric Clinics of North America*, *14*, 241–254. doi:10.1016/j.chc.2004.05.002

Rogers, E. (2003). *The diffusion of innovation* (5th ed.). New York: The Free Press.

Rosenthal, M. B. (2008). Beyond pay for performance—emerging models of provider-payment reform. *New England Journal of Medicine*, *359*, 1197–1200. doi:10.1056/NEJMp0804658.

Shimokawa, K., Lambert, M. J., & Smart, D. (2010). Enhancing treatment outcome of patients at risk of treatment failure: Meta-analytic and mega-analytic review of a psychotherapy quality assurance system. *Journal of Consulting & Clinical Psychology*, *78*, 298–311. doi:10.1037/a0019247.

Slade, K., Lambert, M. J., Harmon, S. C., Smart, D. W., & Bailey, R. (2008). Improving psychotherapy outcome: The use of immediate electronic feedback and revised clinical support tools. *Clinical Psychology & Psychotherapy*, *15*, 287–303. doi:10.1002/cpp.594

Slade, M., Thornicroft, G., & Glover, G.S.O. (1999). The feasibility of routine outcome measures in mental health. *Social Psychiatry and Epidemiology*, *34*, 243–249. doi:10.1007/s001270050139

Strauss, B., Lutz, W., Steffanowski, A., Wittmann, W., Böhnke, J.R., Rubel, J., … Kirchmann, H. (in press). Benefits and challenges in practice-oriented psychotherapy research in Germany: The TK and the QS-PSY-BAY projects of quality assurance in outpatient psychotherapy. *Psychotherapy Research*.

Stricker, G., & Trierweiler, S. J. (1995). The local clinical scientist: A bridge between science and practice. *American Psychologist*, *50*, 995–1002. doi:10.1037/0003-066X.50.12.995

Strupp, H. H., Horowitz, L. M., & Lambert, M. J. (1997). *Measuring patient changes in mood, anxiety, and personality disorders: Toward a core battery.* Washington DC: American Psychological Association Press.

Trauer, T., Gill, L., Pedwell, G., & Slattery, P. (2006). Routine outcome measurement in public mental health—what do clinicians think? *Australian Health Review*, *30*, 144–147. doi:10.1071/AH060144.

Tremblay, M. C., Hevner, A. R., & Berndt, D. J. (2012). Design of an information volatility measure for health care decision making. *Decision Support Systems*, *52*, 331–341. doi:10.1016/j.dss.2011.08.009.

Walfish, S., McAlister, B., O'Donnell, P., & Lambert, M. J. (2012). An investigation of self-assessment bias in mental health providers. *Psychological Reports*, *110*, 639–644. doi:10.2466/02.07.17

Warren, J. S., Nelson, P. L., Mondragon, S. A., Baldwin, S. A., & Burlingame, G.M. (2010). Youth psychotherapy change trajectories & outcome in usual care: Community mental health versus managed care. *Journal of Clinical & Consulting Psychology*, *78*, 144–155. doi:10.1037/a0018544

Weissman, N. W., Allison, J. J., Kiefe, C. I., Farmer, R., Weaver, M.T., Williams, O.D., … Baker, S. (1999). Achievable benchmarks of care: The ABCs of benchmarking. *Journal of Evaluation in Clinical Practice*, *5*, 269–281. doi:10.1046/j.1365-2753.1999.00203.x

West, J. C., Mościcki, E. K., Duffy, F. F., Wilk, J.E., Countis, L., Narrow, W.E., & Regier, D.A. (in press). APIRE Practice Research Network: Accomplishments, challenges, and lessons learned. *Psychotherapy Research*.

Whipple, J. L., Lambert, M. J., Vermeersch, D. A., Smart, D. W., Nielsen, S.L., & Hawkins, E.J. (2003). Improving the effects of psychotherapy: The use of early identification of treatment failure and problem solving strategies in routine practice. *Journal of Counseling Psychology*, *58*, 59–68. doi:10.1037/0022-0167.50.1.59

Wierzbicki, M. & Pekarik, G. (1993). A meta-analysis of psychotherapy dropout. *Professional Psychology: Research and Practice*, *24*, 190–195. doi:10.1037/0735-7028.24.2.190.

Youn, S.J., Kraus, D. R., & Castonguay, L.G. (2012). The Treatment Outcome Package: Facilitating practice and clinically relevant research. *Psychotherapy*, *49*, 115–122. doi:10.1037/a0027932

Zimmerman, M., & McGlinchey, J. B. (2008). Why don't psychiatrists use scales to measure outcome when treating depressed patients? *Journal of Clinical Psychiatry*, *69*, 1916–1919.

Downloaded by [Curtin University Library] at 02:16 29 July 2013