A chance for change : building an outcome monitoring feedback system for outpatient mental health care

Jong, K. de

Chapter 1

General Introduction

Background

Over the last decades critique on the generalizability of scientific studies on the outcome of psychological interventions (e.g. Essock, Drake, Frank, & McGuire, 2003; Westen, 2005; Stirman, DeRubeis, Crits-Christoph, & Brody, 2003) has led to the emergence of practice-based research within the scientific community (Margison et al., 2000). At the same time, there was an increasing demand for practice-based research from therapists and mental health care providers. A combination of factors, such as expanding costs of mental health care and a deteriorating economic situation, has led to political measures, such as a restriction of the amount of refundable therapy sessions by insurance companies (e.g. Lambert, Huefner, & Nace, 1997). This state of affairs has made it necessary for providers to show that their interventions are effective. Measuring outcomes was also stimulated by the quality assurance perspective that has been adopted by the field (e.g. Edmunds et al., 1997; Valenstein et al., 2004).

Outcomes of psychological interventions depend on many factors, including patient and therapist characteristics and treatment variables. Results from randomized controlled trials tend to show positive results of treatment, but results measured in clinical practice are less favourable (Barkham et al., 2008; Hansen, Lambert, & Forman, 2002; Hansen, Lambert, & Forman, 2003; Weisz, Donenberg, Han, & Weiss, 1995). There has been a growing awareness in research and practice that treatment effects should be evaluated for the individual patient in everyday practice and that the treatment has to be adjusted when patients do not progress according to expectation (Lambert et al., 2003). Measuring patient progress in clinical practice is often referred to as (routine) outcome monitoring or measuring. In outcome monitoring therapists are typically provided with feedback on their patients' progress using a generic outcome measure. In some models, the progress of the patient is benchmarked against a prediction model (Lambert, 2007). Recently, outcome monitoring has burgeoned and there is a variety of feedback models available, that not only differ in design, but also in effectiveness. Recent review articles and meta-analyses show that feedback can have slightly negative to very large positive effects (Carlier et al., 2010; Knaup, Koesters, Schoefer, Becker, & Puschner, 2009; Shimokawa, Lambert, & Smart, 2010). In addition to therapists being able to adjust the treatment of individual patients, outcome monitoring enables research on patient, therapist and treatment factors that may be predictive of positive outcomes, which may inform decision making on an aggregated level (e.g. treatment unit, organization, national or international level).

At the start of this project, Michael Lambert was one of the few who had done controlled research in this area. His group still has performed the largest number of studies, but it was uncertain whether results of these studies could be generalized to Dutch outpatients, since his research was carried out in a very specific setting. The

majority of the patients in his studies were students at a private mormon (Latter Day Saints; LDS) university, who were seeking counseling in the university counseling center, resulting in highly educated, young patients of relatively high social economic status (Shimokawa, et al., 2010). In addition, practically all of the therapists are LDS church members as well. This resulted in a very specific cultural subgroup of patients and therapists and the results of his studies might not translate well to other American cultural groups, let alone Dutch patients and therapists. Apart from cultural differences, the American mental health care system is different from the system in the Netherlands, where almost all mental health care is in the public sector and even service provided by private practitioners is reimbursed by public funds.

The central objective of this thesis is to develop an effective outcome monitoring feedback model that is fit for outpatient mental health treatment in the Netherlands and that may improve treatment outcomes for patients. The purpose of this chapter is to give an overview of the relevant literature. The chapter is organized as follows. First, an overview of relevant literature on outcome monitoring is given. Second, a definition of outcome, treatment success and treatment failure is provided. Third, relevant predictors of outcome will be discussed. Finally, the need for feedback to therapists is discussed. The chapter ends with an overview of this thesis.

Outcome monitoring

As has been stated above, outcome monitoring refers to the (frequent) measuring of results during therapy. The measures that are used depend on the treatment goals, which may differ distinctly across treatment settings. Worldwide, there are many large initiatives in which outcomes are monitored in clinical practice, including the COMPASS (Howard, Moras, Brill, Martinovich, & Lutz, 1996), the Partners for Change Outcome Management System (Miller, Duncan, Sorrell, & Brown, 2005), and the Treatment Outcome Package (Kraus, Seligman, & Jordan, 2005) in the United States; the Clinical Outcomes in Routine Evaluation – Outcome Measure (Evans et al., 2002) in the United Kingdom; and the Health of the Nation Outcome Scales (HoNOS) in the United Kingdom, Australia and New Zealand (Burgess, Pirkis, & Coombs, 2006; Wing et al., 1998).

In the Netherlands, a large national project on routine outcome monitoring started in 2009. The project is a collaboration between the branch organization of mental health care providers, the branch organization of health insurance companies, the national patient platform for mental health care, the Dutch institute for psychologists, the Dutch organization for psychiatrists and the national organization for patients in mental health care. It aims to promote routine outcome monitoring in clinical practice for all patient groups (Stuurgroep ROM ggz, 2010). What is unique about the project is that it is not centered around the use of one instrument or patient group, but a wide range

of instruments and patients. There are several task forces in the project that advice the field on topics such as the use of instruments, measurement frequency, comparability of outcome measured by different instruments, implementation and infrastructure.

Outcome monitoring has been promoted by several professional organizations. The American Psychological Association (APA) Presidential Task Force on Evidence-Based Practice in Psychology (2006) states that "providing clinicians with real-time patient feedback to benchmark progress in treatment and clinical support tools to adjust treatment as needed" is one of "the most pressing current research needs" (p.278). The APA Interdivisional (Divisions 12 and 29) Task Force on Evidence-based Therapy Relationships (2011) reports that collecting client feedback is demonstrably effective and states in their recommendations for practice: "Practitioners are encouraged to routinely monitor patients' responses to the therapy relationship and ongoing treatment. Such monitoring leads to increased opportunities to reestablish collaboration, improve the relationship, modify technical strategies, and avoid premature termination". Similar advice is given by the American Group Psychotherapy Association (2007) in their Practice Guidelines for Group Psychotherapy.

Outcome monitoring can have several functions. The ROM ggz project mentions four functions of outcome monitoring: 1) to support clinicians in their treatment process; 2) to learn from aggregated outcomes within the organization and compare oneself with other organizations; 3) to demonstrate the effectiveness of treatments to third-party payers, such as insurance companies and the Ministry of Health; 4) to study predictors and mediators of outcome using large national databases (Stuurgroep ROM ggz, 2010). Kazdin (2008) states that the key argument for systematic evaluation is to provide high-quality care. Whether clinicians use evidence-based treatment or individualized treatment, they can never be sure that the treatment will be effective. In addition, monitoring treatment effects in an ongoing way is important to make decisions on continuing, altering or terminating treatment on the basis of how well the patient is doing (Kazdin, 2008). Sapyta, Riemer and Bickman (2005) use the metaphor of learning archery and indicate that without feedback on where the arrow lands, it is impossible to master archery, no matter how much natural talent someone has at archery. They state that: "Without direct feedback on how their clients are progressing, clinicians are essentially wearing a blindfold while shooting at a target". Although the archery metaphor is certainly appealing, in real life the analogy does not hold. In archery feedback is straightforward: you hit the target or not. In clinical practice, the feedback will not always be that clear. When two measurements are administered to a patient, how does the clinician know if he or she is on target or not? How much does a patient need to improve to be considered a successful case? These questions will be addressed in the next paragraph.

Defining outcomes

Treatment failure

One of the main objectives of feedback models in clinical practice is to prevent treatment failure. Until recently, treatment failure was an issue that was not often addressed in literature, but a renewed interest in the matter has emerged, resulting in special issues in *Cognitive Behavioral Practice* (Dimidjian & Hollon, 2011) and the *Journal of Clinical Psychology* (Lampropoulos, 2011), as well as a series of articles in the *American Psychologist* (Barlow, 2010; Castonguay, Boswell, Constantino, Goldfried, & Hill, 2010; Dimidjian & Hollon, 2010). Treatment failure is a complex concept to define, as it may include dropout or premature termination, nonresponse (no change), partial change, slow change, deterioration (negative change) and relapse (failure to maintain gains). In addition, there might be challenges in defining failure in therapy depending on whose perspective is being used: patient, therapist or other observers (Lampropoulos, 2011). It seems that definition of treatment failure also differs between different treatment modalities. A recent special issue of the *Journal of Clinical Psychology* (November 2011) showed treatment failure from the perspective of five different theoretical orientations. In psychodynamic therapy gaining insight and character change are considered the most important goals and treatment could be perceived as failed by psychodynamic therapists if symptomatic relief would occur without structural changes in character (Gold & Stricker, 2011). The behavioral therapy approach focuses on nonresponse and deterioration, defined by symptom attenuation and uses self-report measures to assesses outcomes (Hopko, Magidson, & Lejuez, 2011). Watson (2011), who is an experiential psychotherapist, stresses that for patients who are nonresponders or who are slow to respond, the question of whether treatment has failed is not as easy to answer. In interpersonal therapy, failure includes partial response, nonresponse, worsening and premature termination (Ravitz, McBride, & Maunder, 2011). Common denominators that capture most perspectives on treatment failure across theoretical orientations include nonresponse and deterioration (Lambert, 2011), which is the definition we use in this thesis.

Clinical significance

Nonresponse and deterioration are most often defined by the concept of clinical significant and reliable change (Jacobson & Truax, 1991). These authors define clinical significance as returning to normal functioning and this approach is nowadays the leading method in outcome research. The criterion is twofold: (a) the magnitude of change has to be statistically reliable and (b) by the end of treatment patients have

to end up in a (score) range that renders them either indistinguishable from well-functioning people, or has them in the lower range of dysfunctioning (Jacobson, Roberts, Berns, & McGlinchey, 1999). A cut-off point for normal functioning and a reliable change index are calculated. The method works best when adequate norms are available for both the dysfunctional and the normal population. Three possible ways to calculate the cut-off point for normal functioning are given. The first method uses a score of two standard deviations above the patient population mean as cutoff. The second method is similar to the first and uses a score of two standard deviations below the normal population mean as a cut-off point. These two methods can only be used when the population curves of the normal population and patient population do not overlap. The third method is the most frequently used method and places the cut-off point at the intersection of the dysfunctional and normal population curves. It is estimated by using the mean and standard deviation of the normative samples for both populations. The reliable change index (RCI) is the minimum amount of change that has to occur to be statically reliable. It is usually expressed as the amount of points on a certain measurement instrument that a patient has to improve between pre- and post-treatment measurements. The RCI depends on the reliability of the measurement instrument and the variability of scores.

Within this system a patient is classified as 'recovered' if reliable change has occurred and the cutoff point for clinical significance is crossed. If the cutoff point is not crossed, but reliable change took place, a patient is classified as 'improved'. If reliable change occurs in opposite direction, the patient is classified as 'deteriorated'. Patients are considered nonresponders if they do not improve nor deteriorate (Jacobson, Follette, & Revenstorf, 1984; Jacobson, et al., 1999; Jacobson & Truax, 1991).

There are some limitations to defining treatment success in this way. If the patients end up in the normal range at the end of treatment, but the magnitude of change is not reliable, patients are considered nonresponders, even though they are functioning comparable to non-patients. In addition, if a patient starts treatment in the normal range, no clinical significance can occur. This is especially problematic in naturalistic studies where selection on severity of problems is less common than in controlled trials. On the other end of the spectrum, there are patients who are not likely to ever return to normal functioning and for whom clinical significant change might not be the best criterion for treatment success. Nevertheless, for the majority of patients in outpatient mental health care it will be possible to classify the outcome of therapy according to this method. Despite its drawbacks, the method of clinical significance by Jacobson and colleagues remains the most frequently applied method and will be applied throughout this thesis.

Outcome: rate of change and end state functioning

The method of determining clinical significance provides us with information on how to interpret the result of treatment when therapy has ended, but does not describe the way in which patients change between the beginning and end of therapy. Recently, there is a renewed interest among psychotherapy researchers in studying not only whether treatments work, but also for whom and under what conditions they work, how they work, and why they work (Laurenceau, Hayes, & Feldman, 2007). Outcome monitoring provides data on the course of therapy and it makes sense to analyse all available data rather than just begin and end state functioning. In fact, one of the major problems in measuring outcomes in psychotherapy research is that it can be a challenge to collect data on treatment outcomes at the end of treatment, due to reasons such as drop-out from therapy or research. This difficulty is especially relevant for practice-based studies, with large numbers of patients and reduced control over the measurements by the researchers.

The availability of large naturalistic data sets that are the result of practice based studies has fired an interest in longitudinal models that are more flexible in handling unbalanced longitudinal data, including survival analysis and multilevel analysis (Singer & Willet, 2003). These models tend to describe change over time and predictor variables are modelled as interaction with time, meaning that they predict the speed of recovery – referred to as the rate of change – rather than end state functioning. To illustrate the difference between outcome defined as end state function and rate of change, consider the following example: two patients can have the exact same score at the start of therapy and at the end of therapy, so the absolute change they made is equal. However, Patient 1 received 10 sessions of therapy, whereas Patient 2 needed 30 sessions. Both have the same end state functioning, but the rate of change is higher in Patient 1. Alternatively, two other patients could have the same starting point and the same rate of change but different treatment durations, resulting in differences in end state functioning. In this thesis, both end state functioning and rate of change will be used as outcome variables.

Predictors of outcome

A vast amount of literature exists on the subject of predicting treatment outcomes. It would be beyond the scope of this thesis to give a full review of all the research that is done in this area, since sufficient review literature is available (e.g. Beutler et al., 2004; Clarkin & Levy, 2004; Lambert & Ogles, 2004; Orlinsky, Ronnestad, & Willutzki, 2004). It is relevant to note that there are two lines of research that look into predictor variables: one focusing on treatment-specific effects and one focusing on common

or non-specific factors. Lutz (2002) states that identifying particular treatments for particular diagnoses could be seen as a bottom-up approach, whereas the top-down approach assesses a class of treatments defined by overlapping techniques, mechanisms and proposed outcomes and could be seen as a first step for large-scale outcome monitoring. Both are relevant for the field of outcome research. For the current research we are interested in the full range of patients that are in outpatient mental health care and not in specific groups of patients. Therefore we take the top-down approach and look for common factors that predict treatment outcomes.

Predictor variables with a demonstrated relationship with treatment outcome can be classified into three categories: patient characteristics, therapist characteristics and treatment characteristics. The aim of the prediction model is to predict the treatment course and end state functioning of new patients; therefore, the focus is on pre- and early treatment predictors. that can be identified at the beginning of treatment, to make early prediction possible. Patient characteristics include both sociodemographic variables (e.g. sex, age, marital status) and clinical variables (e.g. diagnosis, prior treatment). The influence of demographic variables on outcome is mixed and inconsistent (Clarkin & Levy, 2004). Clinical variables have been studied intensively and frequently found predictors include initial severity of symptoms, having multiple Axis I or Axis II diagnoses, or a combination of Axis I and II disorders, prior treatment and duration of psychiatric complaints.

A second category of outcome predictors are characteristics of the treatment received by the patient. These characteristics include a wide range of variables associated with therapeutic techniques, the duration and frequency of therapy sessions, the treatment modality (e.g. individual, group, couples counselling, etc) and interactions between the patient and therapist. It is beyond the scope of this thesis to discuss the full range of treatment predictors. In naturalistic studies, treatment variables usually are outside the control of the researcher and treatments are often modified during the treatment process, making treatment characteristics complicated to use as predictors. Variables that are of interest to us are two early treatment factors that are consistently associated with outcomes: expectancies and the working alliance. Although expectancies could be considered patient characteristics, they usually emerge in the interaction with the therapist and can be modified within that interaction (Westra, Constantino, & Aviram, 2011). Expectancies have a small positive effect on outcome (Constantino, Arnkoff, Glass, Ametrano, & Smith, 2011). The relationship between therapist and patient, referred to as the working alliance or therapeutic alliance, has been the focus of many studies and appears to be a consistent albeit moderate predictor of treatment outcome (Hentschel, 2005; Horvath & Symonds, 1991; Martin, Garske, & Davis, 2000).

Numerous studies have demonstrated that therapist factors explain part of the

variance in outcomes between patients (Crits-Christoph et al., 1991; Dinger, Strack, Leichsenring, Wilmers, & Schauenburg, 2008). Literature shows that therapist effects are usually small in randomized controlled trials and small to medium in naturalistic studies (Crits-Christoph & Gallop, 2006). Little is known about variables that predict therapist effects: demographic characteristics fail to emerge as predictors of outcome, but variables that are more closely related to personality traits for the therapist have been found as predictors, although more research is needed on the subject (Anderson, Ogles, Patterson, Lambert, & Vermeersch, 2009; Beutler, et al., 2004). In order to predict the treatment course and end state functioning for new patients, therapist variables may not be of much use, since it is still quite unclear how they affect outcomes. However, what is interesting for this thesis is how therapist variables may interact with the use of feedback by the therapist. Therefore, therapist variables that have found to be related to acceptance and use of feedback will be included in the study.

Feedback to therapists

The need for feedback

One could wonder why therapists need feedback on their patients' progress. In fact, many therapists claim that they do not need feedback and are perfectly able to monitor progress in their patients themselves. However, meta-analysis comparing 67 studies showed that statistical prediction had somewhat greater accuracy than clinical judgement (Ægisdóttir et al., 2006). In his review of the literature on clinical decision making, Garb (2005) concludes that psychologists should reduce their reliance on informal observation and clinical validation when making clinical judgements or interpret tests. None of these studies have been done on decisions on patient progress. In a study more specifically tailored to recognizing patients that deteriorate, Hannan and colleagues (Hannan et al., 2005) asked therapists to rate their entire caseload on the risk of deterioration for three consecutive sessions. Therapist were able to identify less than 1% of patients that deteriorated. The statistical model, using expected treatment recovery curves, was able to correctly identify 85% of all patients that deteriorated. More recently, Hatfield and colleagues showed that therapists had considerable trouble in identifying deteriorating clients when using their case files. Although therapists did seem to recognize the factors that might be related to deterioration, such as symptom increase and therapy indications such as ruptures in the working alliance or treatment goal failure, they still had trouble noticing these cues in their patients (Hatfield, McCullough, Frantz, & Krieger, 2010).

That clinicians may need feedback on their patients in order to prevent negative outcomes, does not mean that they like it. Some clinicians do not want to use outcome

monitoring feedback, even if it is beneficial to the patient (Aoun, Pennebaker, & Janca, 2002; Walter, Cleary, & Rey, 1998). Feedback theory provides several factors that might be related to the attitude that therapists have towards feedback and whether they will use it. Clinicians will be more likely to use feedback effectively if they perceive the feedback as valid and reliable, pay attention to the feedback, have a preference for externally generated feedback and are committed to use the feedback in their treatment (Claiborn & Goodyear, 2005; Herold & Fedor, 2003; Riemer & Bickman, 2011).

Towards building an outcome monitoring feedback model

As was mentioned before, the main objective of this thesis was to develop an outcome monitoring feedback model for outpatients in the Netherlands and test whether providing feedback on patient progress to therapists can improve treatment outcomes. Therefore, the first step in the process of building a feedback model was to get an idea on how Dutch patients score on the outcome measure we wanted to use – the Outcome Questionnaire – 45 (OQ-45; Lambert et al., 2004) (see Chapter 2) and to obtain more insight in the progress in therapy that is made by these patients over time. In addition, we were interested in the pre and early treatment variables that predict outcomes and the differences among predictive variables in predicting the rate of change and end state functioning (Chapter 4).

In designing a study to measure the effect of feedback, we needed to take into account that the feedback about the patient's progress is provided to therapists (and sometimes the patients) and therapists might differ in how they handle the feedback, which in turn might influence the effectiveness of the feedback. In the study design, and more specially in estimating the amount of patients that needed to be included in the study to have sufficient statistical power, we needed to take into account that the data have a hierarchical structure with three levels: patients were nested within therapists and measurements were nested within patients. In Chapter 3 the factors that influence study design and study power in a situation with therapists accounting for some of the variance of the outcomes in patients is described.

Finally, two feedback studies were conducted. The first study was a multicenter study in an outpatient mental health care setting and followed patients in their treatment for one year. Feedback was provided to the therapist only and in addition to patient outcomes, and therapist characteristics on relevant traits that are hypothetically related to feedback effectiveness were studied (see Chapter 5). The second study was conducted in both public outpatient centers and in private practices and contained both short and longer term therapies, up until two years. This study had three treatment conditions, a control group, a group with feedback to therapists alone and a group in which both therapists and patients received feedback about the patient's

progress (see Chapter 6). Finally, in the general discussion, results of the studies will be summarized and discussed, study strengths and limitations will be reviewed and implications for future research and clinical practice will be presented (see Chapter 7).