RESEARCH PROTOCOL

PROTOCOL TITLE 'Taking feedback to the next level: Efficacy of expected treatment response feedback to therapists

Protocol ID	NL30987.058.09
Short title	Efficacy of ETR feedback
Version	3
Date	September, 29, 2010
Coordinating investigator/project leader	Kim de Jong, Msc GGZ Noord-Holland-Noord, Research Department P.O. Box 18 1850 BA Heiloo Tel: +31 72 5312788 Fax: +31 72 5312660 e-mail: k.dejong@ggz-nhn.nl
Principal investigator(s) (in Dutch: hoofdonderzoeker/uitvoerder) ***Multicenter research: per site***	Kim de Jong (see contact information above) Kim la Croix, MSc Dimence Locatie Rielerenk Nico Bolkesteinlaan 65 7416 SE Deventer Tel: +31 570 63 92 40 Fax: +31 570 63 92 99 e-mail: k.lacroix@dimence.nl
Sponsor (in Dutch: verrichter/opdrachtgever)	*Not applicable (self-sponsored study by participating GGZ-institutions)*
Independent physician(s)	Victor Buwalda, MD Altrecht Lange Nieuwstraat 119 3512 PG Utrecht Tel: +31 6 45106236 E-mail: imsep@xs4all.nl

PROTOCOL SIGNATURE SHEET

Name	Signature	Date
For non-commercial research, **Head of Department:** ***Dr. M.A. Nugter***		21-02-2010
Principal Investigator: ***Drs. Kim de Jong***		21-02-2010

TABLE OF CONTENTS

1. INTRODUCTION AND RATIONALE 8
2. OBJECTIVES 10
3. STUDY DESIGN 10
4. STUDY POPULATION 10
4.1 Population (base) 10
4.2 Inclusion criteria 11
4.3 Exclusion criteria 11
4.4 Sample size calculation 11
5. TREATMENT OF SUBJECTS 12
5.1 Investigational product/treatment 12
5.2 Use of co-intervention 12
6. METHODS 12
6.1 Study parameters/endpoints 13
6.1.1 Main study parameter/endpoint 13
6.1.2 Secondary study parameters/endpoints 13
6.1.3 Other study parameters 13
6.2 Randomisation, blinding and treatment allocation 13
6.3 Study procedures 13
6.4 Withdrawal of individual subjects 16
6.5 Replacement of individual subjects after withdrawal 16
6.6 Follow-up of subjects withdrawn from treatment 16
6.7 Premature termination of the study 16
7. SAFETY REPORTING 17
7.1 Section 10 WMO event 17
7.2 Adverse and serious adverse events 17
7.2.1 Annual safety report 17
7.3 Follow-up of adverse events 18
8. STATISTICAL ANALYSIS 18
8.1 Descriptive statistics 18
8.2 Multivariate analysis 18
8.3 Interim analysis 18
9. ETHICAL CONSIDERATIONS 19
9.1 Regulation statement 19
9.2 Recruitment and consent 19
9.3 Benefits and risks assessment 19
10. ADMINISTRATIVE ASPECTS AND PUBLICATION 19
10.1 Handling and storage of data and documents 19
10.2 Amendments 20
10.3 Annual progress report 21
10.4 End of study report 21
10.5 Public disclosure and publication policy 21
11. REFERENCES 22

LIST OF ABBREVIATIONS AND RELEVANT DEFINITIONS

ABR	ABR form (General Assessment and Registration form) is the application form that is required for submission to the accredited Ethics Committee (ABR = Algemene Beoordeling en Registratie)
AE	Adverse Event
AR	Adverse Reaction
CA	Competent Authority
CCMO	Central Committee on Research Involving Human Subjects
CV	Curriculum Vitae
DSMB	Data Safety Monitoring Board
EU	European Union
EudraCT	European drug regulatory affairs Clinical Trials GCP Good Clinical Practice
IB	Investigator's Brochure
IC	Informed Consent
IMP	Investigational Medicinal Product
IMPD	Investigational Medicinal Product Dossier
METC	Medical research ethics committee (MREC); in Dutch: medisch ethische toetsing commissie (METC)
(S)AE	Serious Adverse Event
SPC	Summary of Product Characteristics (in Dutch: officiële productinfomatie IB1-tekst)
Sponsor	The sponsor is the party that commissions the organisation or performance of the research, for example a pharmaceutical company, academic hospital, scientific organisation or investigator. A party that provides funding for a study but does not commission it is not regarded as the sponsor, but referred to as a subsidising party.
SUSAR	Suspected Unexpected Serious Adverse Reaction
Wbp	Personal Data Protection Act (in Dutch: Wet Bescherming Persoonsgevens)
WMO	Medical Research Involving Human Subjects Act (Wet Medisch-wetenschappelijk Onderzoek met Mensen

SUMMARY

Rationale: Nowadays, many mental health care organizations measure their patient's progress through routine outcome monitoring. The way that progress is measured and the way therapists are provided with feedback on their patient's progress strongly differs between organizations. The question is whether all methods are as effective, and there is surprisingly little research on the effectiveness of routine outcome monitoring available.
Research in the United States shows that providing feedback to therapists based on a prediction model, can improve patient outcome, especially for those patients that are not progressing well in treatment (o.a. Lambert, 2007). The prediction model that was used in the studies by Lambert was calculated based on the initial severity in patient disfunctioning, measured with the Outcome Questionnaire (OQ). Lambert and colleagues conducted five randomized controlled trials on the effectiveness of feedback based on the prediction model en found that in the 'not on track' group (patients that were not progressing well) treatment outcomes were significantly better in the experimental group than in the no feedback control condition. In the 'on track' group, the feedback did not have an effect on outcome.
In the Netherland, a prediction model for Dutch patients was predicted, based on almost 2000 patients in four mental health care organizations. Our model uses the initial severity of patients disfunction, as well as patients' expectancies on treatment outcome as predictors for progress. Results showed that this was also a significant predictor for progess in the data that was collected.
Objective: In this study, two forms of feedback and one control condition are compared:
* Control group: the therapist gets no feedback.
* Outcome monitoring feedback: the therapist gets feedback on the patients progress in a progress chart.
* Complex feedback based on the Dutch prediction model: the therapist gets a progress chart that compares the patient's actual progress with the expected treatment response.
Main questions:
1. Does providing feedback to therapists improve treatment outcome?
2. Does providing feedback based on the prediction model lead to better outcomes than progress feedback alone?
Secundary questions:
1. How well can therapists predict their patient's progress?
2. Are there differences between therapists in treatment outcomes and is this related to the way they use the feedback?
Study design: A 2-year randomized controlled clinical trial on the effectiveness of feedback interventions for therapists
Study population: The study population consists of psychiatric patients between 18 and 65 years that apply for outpatient individual psychiatric treatment in one of the three participating treatment facilities. Patients with psychotic disorders, a current crisis, a high risk of decompensation, a current severe manic episode or insufficient mastery of the Dutch language in reading and speaking are excluded from participating in the study. Patients can have supplement treatments beside the individual treatment and stay in the study, as long as

the individual treatment remains the main treatment. Patients are excluded from the study if individual treatment is not their main treatment.
Intervention (if applicable): The intervention consists of providing feedback to therapists on their patients' progress. In the outcome monitoring feedback condition the therapist gets feedback on patient progress in progress charts and tables. The progress of the patient can be viewed by the therapist at all times, by logging on to the feedback system (RequestXL), but is also actively provided by e-mail at session 1, 3, 5, 10 and 15.
In the complex feedback condition with prediction model the actual treatment course (based on the OQ scores) of the patient is compared with the predicted treatment course. The expected treatment course is calculated by a formula. The progress of the patient can be viewed by the therapist at all times, by logging on to the feedback system (RequestXL), but is also actively provided by e-mail when the patient is not progressing well. The therapist then receives an e-mail with high urgency.
In the complex feedback condition feedback is also provided on the ASC. The ASC is administered when the patient goes off track (through the 75% negative bound of the confidence interval around the predicted treatment course) and measures the therapeutic alliance, motivation, social support and life events. The ASC is combined with so called Clinical Support Tools, a set of Microsoft Word documents that provide practical tips on improving the therapeutic alliance, motivation and social support. The practical tips are based on a literature review on these topics.
Main study parameters/endpoints: The primary outcome measure of the study is patients' disfunctioning on the Outcome Questionnaire. Mores specifically, the progress that the patient makes during treatment. The results will be analysed using multilevel analysis, which has the advantage of being able to handle missing data really well. Therefore, data from all patients, including those that dropped out of treatment or dropped out of the study can be used in analysis.
Nature and extent of the burden and risks associated with participation, benefit and group relatedness:
Burden: The burden for the patient consists of completing a 5 minute questionnaire before each treatment session, for a maximum of 15 times. After the research period (end of treatment or after 15 session) they will get two follow-up measures after 3 and 6 months. The average treatment duration for outpatient indivual treatment was around 9 sessions in a previous study at one of treatment settings. If patients go off track, they are asked to complete an addition questionnaire, that also takes about 5 minutes to complete.
In addition to the self-report questionnaires, patients are getting a diagnostic interview that takes approximately 1 hour.
The burden for the therapist consist of completing brief questionnaires during the treatment of a participating patient (total burden around 10 minutes) and completing a longer questionnaire once every 6 months (about 20 minutes per adminstration, total of 5 administrations).
Risks: The risks for both the patient and the therapist are minimal.

1. INTRODUCTION AND RATIONALE

A large body of research, performed over 40 years, has demonstrated the general effectiveness of psychological treatments (Lambert & Ogles, 2004). However, while treatment is beneficial for many patients, not all patients profit from therapy. In randomized controlled clinical trials, where treatment is provided under more or less optimal conditions, an average of 67% of patients is statistically reliably improved at the end of treatment. In clinical practice, the success rates are much lower: in an American national database of more than 6000 patients, only 35% of the patients was improved (Hansen, Lambert, & Forman, 2002).

Riemer & Bickman (2003) consider feedback interventions a promising approach to improve clinical practice. According to them, clinicians need to have more systematic and reliable information about the status of their patients, so that they are capable of adjusting their treatment. Literature from psychology, educational science and medicine shows the effectiveness of feedback interventions (e.g. Kluger & DeNisi, 1996, 1998; Jamtvedt et al., 2006; Hattie & Timperley, 2007; Veloski et al., 2006). A general meta-analysis by Kluger & DeNisi (1996) showed an average effect size of 0.38 for feedback interventions. More specifically tailored to the effectiveness of health status feedback to professionals, a meta-analysis by Sapyta (2004, in Sapyta, Riemer, & Bickman, 2005) found an average effect size of 0.21. Further analysis showed that feedback on 'flagged patients' (those not progressing well in treatment) had an average effect size of 0.31, while there was no significant feedback effect for the 'non-flagged' samples.

In psychotherapy the research group of Michael Lambert has published five controlled studies in which therapists received feedback about a patient's improvement through the use of progress charts and warnings about patients who were not demonstrating the expected treatment response (Hawkins, Lambert, Vermeersch, Slade, & Tuttle, 2004; Lambert, et al., 2002; Lambert, et al., 2001; Whipple, et al., 2003)Harmon et al., 2007. Results showed that patients at risk for treatment failure in the feedback condition showed significantly more improvement than at risk-patients in the no-feedback control condition. This effect was even stronger when feedback was combined with a clinical support tool. The effect sizes for the difference between patients at risk in the various feedback conditions and control conditions ranged from 0.34 to 0.92. Feedback had no significant effect in the non-flagged samples (Lambert, 2007). Other research groups have focused mainly on calculating an expected treatment response to signal patients at risk for treatment failure (Lueger, et al., 2001; Lutz, et al., 2006)(e.g. Lutz, Martinovich & Howard, 1999), but in psychotherapy Lambert and colleagues are the only ones that have studied the effect of feedback in a controlled design.

In order to get an idea as to why providing therapist with feedback works, Hannan et al (2005) performed a study in which therapists were asked to rate their patients' progress and predict patient outcome in a period of three weeks. The therapist's predictions were compared with a statistical prediction model and with actual outcome. Therapists tended to overestimate the percentage of positive outcomes and underestimate the percentage of patients that did not change or deteriorated during treatment. In 332 patients, 26 patients actually deteriorated, of which the statistical procedure identified 77% correctly, while

therapists identified only 0.04% of the patients correctly as a treatment failure. This study shows that therapists might not be very good in predicting their patients' progress, when it comes to predicting poor outcomes. By contrast , in a randomized controlled trial comparing two treatments for borderline personality disorders Spinhoven, Giesen-Bloo, Van Dyck & Arntz (2008) found that therapists and assessors could accurately predict indices of outcome above and independent of patient characteristics. Surprisingly few studies have been performed in this area, but the two available studies seem to show mixed results.

The studies performed by the research group of Lambert are not without limitations. Most of the studies have been conducted in the university counseling center, where students are the main treatment population (Lambert, 2007). Only one of the studies has been conducted in a community outpatient setting (Hawkins et al., 2004). Another critique has been that hardly any replications were outside the research group of Lambert; the only one being in a 30-day inpatient treatment in Switzerland (Berking, Orth & Lutz, 2006). Part of the reasons there have been so few replications outside of Lambert's group is that in order to replicate the studies by Lambert, one needs a prediction model to identify the off-track patients. If patients deviated too from the expected treatment response predicted by the model, and exceed the error bound around this prediction, they are considered to be off-track. Developing a prediction model for expected treatment response takes a substantial amount of effort and most research groups have either not invested in that, or have not been able to get enough data yet to develop a model. Fortunately, there is a prediction model available for Dutch mental health care (de Jong et al, in preparation).

One of the issues that is important in feedback studies is whether merely providing the therapist with feedback is enough. According to general theories of feedback, when a person receives feedback a comparison is made between the content of the feedback and a goal or standard (Kluger & DeNisi, 1996). In clinical practice, therapists are usually motivated to achieve improvement or recovery in their patients. When presented with feedback, for instance patient progress charts, a comparison is made between the goal (progress towards recovery) and the feedback (progress so far). The result of this feedback-standard comparison creates a feedback sign – a positive or negative evaluation of one's performance relative to the goal. When a discrepancy is noted, people are motivated to reduce it. Following this line of thinking, it makes sense that feedback does not work in non-flagged case, as there is no discrepancy in these cases, so therapists are not motivated to change their behavior.

The effects of feedback are complex and there are numerous factors that influence the impact of feedback. In feedback studies in clinical practice, especially the factors associated with the recipient of the feedback, the therapist, seem important. If a therapist does not look at the feedback, or does not take it seriously, the feedback intervention might loose its effect. Feedback is more likely to be accepted if it comes from a source that has credibility, a combination of expertise and trustworthiness, and has personal relevance to the receiver (Clairborn & Goodyear, 2005; Sapyta, et al., 2005). Another variable that seems important in whether external feedback will be accepted by the recipient is the feedback preference or propensity of that person. Herold and colleagues (Herold & Fedor, 2003)(Herold, Parsons &

Rensvold, 1996; Herold & Fedor 1998) report that some people have greater preference for externally mediated feedback than others. They refer to this preference as having an *external feedback propensity*. Sapyta et al. (2005) stress that goal commitment is another important recipient variable. Goal commitment is the amount of interest a person has in accomplishing a goal, even if it requires time and effort. Self-regulation theories propose that a discrepancy between a current actual and a desired goal state causes a person to change behavior only if the person is both committed to the goal and aware of the discrepancy. A person is committed to a certain goal mainly if the goal is attractive to him and if he believes it is reasonable for him to accomplish the goal.

2. OBJECTIVES

Primary Objectives:

1) Does providing therapists with feedback about their patients improve patient outcomes?
2) Does providing feedback based on an expected treatment response result in better outcomes than routine outcome monitoring feedback?

Secondary Objectives:

1) Can therapists accurately predict their patient's progress, with or without getting feedback?
2) Are therapists' feedback propensity and goal commitment predictive of better outcomes in the feedback conditions?

3. STUDY DESIGN

A two-year randomized controlled, single-blinded design in three outpatient non-academic settings within two mental health care organizations.

4. STUDY POPULATION

4.1 Population (base)

Patients. Approximately 800 patients between 18 and 65 years, that apply for individual psychotherapy or psychological treatment at one of the three research locations (Schagen, Deventer, Twello), will be asked to participate in the study.
Therapists. Therapists are psychotherapists, psychologists or psychiatric nurses that work at the three research locations. Trainees will be excluded from the study. Approximately 50 therapists will be included in the study.

4.2 Inclusion criteria

Patients that apply for outpatient individual psychiatric treatment in one of the three research locations.

4.3 Exclusion criteria

Exclusion criteria are: having a psychotic disorder as main diagnosis, having a current severe manic episode, a high risk of decompensation, insufficient language skills in Dutch reading and speaking. Patients will be randomly allocated to one of the three research conditions.

4.4 Sample size calculation

Performing an a priori power analyses for multilevel analysis is complicated, because one has to have a good estimate of the variance components in advance and these are usually unknown, especially for three level models and three research conditions (de Jong, Moerbeek & Van der Leeden, 2009). An a priori power analysis was performed for two research conditions, based on the variance components of previously collected data, which resulted in approximately 4 patients per therapist, and 50 therapists participating for a power of 0.80, which is generally considered to be sufficient. However, this analysis was performed with an intended moderate effect (Cohen's *d* = 0.50), whereas feedback meta-analysis shows a small effect (Cohen's *d*=0.30). Also, in the current study we have three research conditions, which is considerably more complicated to model. The difference between the two feedback conditions is most likely smaller than *d*=0.30. Another thing that needs to be considered is that the effect only applies to the "not on track" group. As this group consists of approximately 25% of the total patient group and we cannot identify this group beforehand, the number of patients that need to be included in the study needs to be four times as large as the results from the a priori power analysis.

A traditional a priori power analysis was performed in G*Power 3.0.10 for F tests with repeated measures, three research groups, three repeated measures, a power coefficient of 0.80, an α of 0.05, and a correlation of 0.70 between the repeated measures. This resulted in a needed sample size of 90 not on track patients. The total sample size needed would then be 360 patients. This analysis does not take into account that patients are nested within therapists and therapists will explain part of the variance. We know from a previous study that the intraclass correlation was 0.18 at the therapist level. A rule of thumb to correct for the variance at the therapist level is to use the formula for the design effect: 1 + (*k*-1)**icc*, with k being the number of patient per therapist and the *icc* is the intra-class correlation. In our case, having 4 patients per therapist, we would get a design effect of: 1 + 3 * 0.18 = 1,54. This means that 1.54*360 = 554 patients need to be included in the study. As we expect that some of the data will not be valid, the number of 600 patients was chosen as targeted *N*.

The patient flow in the three research locations is large enough to make that number in the inclusion period, that is set at 6 months. Intermediate power analysis will be performed after six months. Based on the observed variance and estimated power for the

included number of patients, the inclusion period may be prolonged. This will influence study duration.

5. TREATMENT OF SUBJECTS

5.1 Investigational product/treatment

In the ETR feedback condition, the therapist will receive feedback about the patient's progress, bases on an expected treatment response model. The ETR model is based on data collected in three different mental health care organizations in the Netherlands (de Jong et al, in preparation). Therapists receive an alert when the patient deviates from the expected track. An error bound around the expected treatment response for the patient indicates when the therapist is signals. If 75% failure boundary is crossed by the patient, the therapist receives an orange warning signal, indicating that the patient has an increased chance of treatment outcome. If the 95% failure boundary is crossed, the therapist receives a red warning signal, indicating that the patient is most like to have poor treatment outcome if he or she continues on the current track. When the patient signals orange or red, the Assessment for Signal Clients questionnaire will be administered to the patient, and the therapist will receive feedback about the scores on the ASC, in addition to the scores on the OQ-45.
In the ROM feedback condition, feedback is given in the form that most routine outcome monitoring feedback is being provided. The therapist gets a graph of the patient's progress after session 1, 3, 5, 10 and 15. In the graph, the patient's progress on the OQ-45 is monitored. The therapist receives feedback on the OQ total score, the subscale scores and the critical items.
In the control condition, therapists will receive no feedback about their patient's progress. All therapists can be in all three conditions, as randomization takes place at the patient's level.

5.2 Use of co-intervention

Patients are not excluded for getting co-interventions that supplement the individual treatment. Pharmacological treatment, as well as other treatment forms are allowed. However, if individual treatment is no longer the main treatment, patients are excluded from the study (e.g. a patient starts out with individual treatment and is referred to group therapy. He or she might still see the individual therapist, but on a low frequency, the study is stopped for this patient then. Data collected up until then will still be used in analysis) .

6. METHODS

6.1 Study parameters/endpoints

6.1.1 Main study parameter/endpoint

The main study parameter is the level of the patient's dysfunctioning, measured by the Outcome Questionnaire-45 (OQ-45). Data from session-to-session measurements will be analysed using multilevel modelling. A maximum of 15 sessions was chosen to be analysed. The average amount of sessions is around 9 in the participating centers, but ranges from 1 to more than 100 sessions. If patients are measured until the end of treatment, the longer treatments will have too much influence in the model (because they have more measurement points). As a result, the model is no longer representative for brief treatments. Approximately ¾ of treatments are done in 15 sessions, and that is why this number was chosen as a basis for data analysis.

6.1.2 Secondary study parameters/endpoints

Two follow-up measures are included in the study design: one at three months and one at six months after treatment termination or 15 sessions of treatment.
Another secondary study parameter includes the percentage of treatment drop-outs in each reseach condition.

6.1.3 Other study parameters

Other study parameters include initial severity on the OQ-45 at start of treatment, diagnosis on Axis I and II of the DSM-IV, gender, age and previous treatment for patients. Therapist factors that are studied are: gender, age, years of experience, use of the feedback, therapist expectation of outcome, internal and external feedback propensity and therapist's attitude towards using the feedback.

6.2 Randomisation, blinding and treatment allocation

Patients are randomly assigned to one of the three research conditions. The randomization procedure is build in the software used to provide feedback and is completely random. The researcher is aware of the treatment condition of the patients, but cannot influence it. By the nature of the intervention, therapists are aware of the research condition of their patients.

6.3 Study procedures

Patient measurements

Outcome Questionnaire-45 item version (OQ-45). The Dutch translation of the Outcome Questionnaire-45 item version (OQ-45) is used to measure patient progress during treatment. The OQ-45 is a self-report instrument and has 45 items, 9 of which are

reversed, asking how the respondent has felt over the last week on a 5 point rating scale, ranging from 0 (never) to 4 (almost always). The OQ-45 consists of three subscales that are aimed at assessing different domains of client functioning: Symptom Distress, Interpersonal Relations and Social Role. The Symptom Distress domain consists 25 items relating to psychological symptoms that are common in highly prevalent mental disorders. The Interpersonal Relations domain consists of 9 items that assess the functioning of the patient in interpersonal relationships, and the Social Role domain assesses the patients functioning in social roles, such as work and school. The Dutch OQ-45 has appropriate psychometric properties. The internal consistency for the Total score ranges between 0.92 and 0.96 in university, community, patients and community and patients combined samples. For the subscales the international consistency is 0.90-0.95 for the Symptom Distress scale, 0.74-0.84 for the Interpersonal Relations subscale and 0.53-0.72 for the Social Role subscale (de Jong, Nugter, Lambert & Burlingame, 2008). An additional factor that measures Anxiety and Somatic Distress was found, that can be used as separately from the three factor structure. This factor had an internal consistency ranging from 0.79 to 0.89 (de Jong et al, 2007).

Mini International Neuropsychiatric Interview (MINI). The Mini International Neuropsychiatric Interview (MINI) Plus 5.0.0.-R is used to screen for Axis-I disorders. The MINI is a short clinical diagnostic interview developed to explore the presence of Axis-I disorders according to the DSM-IV diagnostic criteria (Sheehan et al., 1998). The MINI-Plus is an extended version of the original MINI. Lecrubier and colleagues (Lecrubier et al., 1997) report sufficient reliability; inter-rater reliability ranged from k = 0.88 to 1.00, test-retest reliability ranged from 0.76 to 0.93, validity was demonstrated by sufficient concordance with the Composite International Diagnostic Interview (CIDI, WHO). In this study the assessment of current psychopathology was used. The Dutch translation of the MINI-Plus 5.0.0-R was used in an electronic version (Van Vliet et al., 2000). The MINI is part of the standard intake procedure on all three research locations.

Assessment for Signaling Clients (ASC). The Assessment for Signaling Clients (ASC) is used to measure the therapeutic alliance, motivation for treatment, social support and stressful life events. The ASC consists of 40 items that are scored on a five point rating scale, ranging from 1 (strongly disagree) to 5 (strongly agree). All subscales have a cutoff score that indicates if that factor is problematic in the patient. In addition, each item has an individual cutoff score indicating when the score on that item is problematic. The Alliance subscale consists of 11 items that assess the alliance in the session with the therapist prior to assessment. An example item is “I felt I could trust my therapist completely”. The Social Support subscale consists of 11 items that assess the support that the patient felt outside of therapy. An example item of this scale is “I could talk about problems with my friends”. The motivation subscale consists of 9 items and assesses the patient’s motivation for therapy. An example item is “I don’t think therapy will help me feel any better”. The Life Events subscale assesses life events in the past week, that may have influenced the patient’s functioning and consists of 9 items. An example item is “I felt rejected or betrayed by someone”. The ASC has been translated into Dutch. A pilot study to assess its psychometric properties is currently being performed.

Demographic questionnaire. The demographic questionnaire is a 19-item self-constructed questionnaire that assesses the demographic characteristics of the patient. It asks for the patient's date of birth, gender, postal area code, nationality, land of birth, land of birth of the patient's parents, marital status, living and working situation, educational level, prior treatment, pretreatment use of medication, the main complaint and the duration of the main complaint.

Therapist measurements

Feedback User Quesionnaire. A feedback user questionnaire will be administered to the therapists every six months and consists of the External Feedback Propensity Scales and an adaptation of the CFIT User Survey. The *External Feedback Propensity Scale* (EFPS; Herold et al, 1996) is being used to measure external feedback propensity. External feedback propensity reflects the preference for externally mediated feedback as well as greater faith in information that one can self-generate. The scale consists of six items that are answered on a five-point rating scale that varies from *strongly disagree* to *strongly agree*. An item from the External Propensity scale is 'It is very important to me to know what people think of my work'. An adaptation of the CFIT User Survey, designed by the Center for Evaluation and Program Improvement of Vanderbilt University, will be used to measure goal commitment, therapy related self-efficacy and perceived validity of the feedback. The items are scored on various five-point rating scales. The goal commitment scale exists of seven items and is based on the Goal Commitment Scale (Hollenbeck & Klein, 1987). A sample item is 'It is hard to take the idea of using these measures in my clinical practice seriously'. The therapy related self-efficacy scale consists of eight items and was constructed according to Bandura's guidelines of how to develop domain-specific self-efficacy scales (Bandura, 1997). A sample item from that scale is 'To what extend do you feel confident in your ability to know what to do if a client is not progressing in treatment'. Perceived validity of the feedback will be measured by a six-item scale, that was self-constructed by Vanderbilt University. A sample item from that scale is: 'I think that feedback based on the OQ-45 will be helpful for my counseling'. In a previous study amongst 47 therapists, internal consistency values were 0.90 for the goal commitment scale, 0.87 for the perceived validity scale and 0.83 for the self-efficacy scale.

Expectation of Outcome Questionnaire (EOQ). A self-contructed questionnaire will be used to measure therapist's expectation of their patient's treatment outcome. This two item questionnaire assesses 1) How much progress the therapist thinks the patient has made since the start of treatment and 2) What treatment outcome the therapist expects for this patient (deterioration, no change, improvement, recovery) at the end of treatment.

Use of Feedback (UOF). A self-constructed questionnaire assesses what the therapist has done with the feedback. The first item assess to what extend the feedback was in concurrence with what the therapist expected. The second item assesses if the therapist discussed the feedback with the patient and the third item assesses what the therapist has done with the feedback as a result of the feedback.

Procedure

All patients that asked to participate in this study will be interviewed with the MINI Plus as part of the standard intake procedure. After patients agree to participate the research assistant will adminster the SCID-II PQ. As from start of treatment, participants will be asked to complete the OQ-45 before each session, for a maximum of 15 sessions and/or one year. In addition, the OQ-45 will be administered 3 and 6 months after treatment (or study) termination. In the ETR feedback condition, patients will fill out the ASC when they are going off track. In the ROM feedback condition and control condition, patients will fill out the ASC when they would have signalled, but the therapist will receive no feedback on how the patient responded.
Therapists will be asked to complete the EOQ on each of their patients at sessions 1, 5, 10 and 15 or the end of treatment, if this is within 15 sessions. The UOF will be administrated after sessions 5, 10 and 15 or the end of treatment. To measure therapists' attitude towards getting feedback, the CFITUS will be administered before the start of the study, after six months, after one year and after a year and a halve. At the first administration, before the start of the study, the EFPS and GSE will be filled out by the therapists as well.

6.4 Withdrawal of individual subjects

Subjects can leave the study at any time for any reason if they wish to do so without any consequences. The investigator can decide to withdraw a subject from the study for urgent medical reasons, or if the patients diagnosis or treatment is changed in such a way that the patient no longer fits the inclusion criteria of the study.

6.5 Replacement of individual subjects after withdrawal

If patients withdraw from the study, the data used up until the moment of withdrawal will be used in analysis. No additional subjects will be included.

6.6 Follow-up of subjects withdrawn from treatment

Patients that withdraw from treatment (drop-outs) will not be followed-up.

6.7 Premature termination of the study

Premature termination of the study is highly unlikely, but in case it will occur (for instance if a reorganization in one of the treatment centers occurs) it will not have unwanted effects for the patient. Patients will be informed of study termination and thanked for their participation.

7. SAFETY REPORTING

7.1 Section 10 WMO event

In accordance to section 10, subsection 1, of the WMO, the investigator will inform the subjects and the reviewing accredited METC if anything occurs, on the basis of which it appears that the disadvantages of participation may be significantly greater than was foreseen in the research proposal. The study will be suspended pending further review by the accredited METC, except insofar as suspension would jeopardise the subjects' health. The investigator will take care that all subjects are kept informed.

7.2 Adverse and serious adverse events

Adverse events are defined as any undesirable experience occurring to a subject during a clinical trial, whether or not considered related to the investigational intervention. All adverse events reported spontaneously by the subject or observed by the investigator or his staff will be recorded.

A serious adverse event is any untoward medical occurrence or effect that at any dose results in death;

- is life threatening (at the time of the event);
- requires hospitalisation or prolongation of existing inpatients' hospitalisation;
- results in persistent or significant disability or incapacity;
- is a congenital anomaly or birth defect;
- is a new event of the trial likely to affect the safety of the subjects, such as an unexpected outcome of an adverse reaction, lack of efficacy of an IMP used for the treatment of a life threatening disease, major safety finding from a newly completed animal study, etc.

All SAEs will be reported to the accredited METC that approved the protocol, according to the requirements of that METC.

7.2.1 Annual safety report

In addition to the expedited reporting of SUSARs, the sponsor will submit, once a year throughout the clinical trial, a safety report to the accredited METC, competent authority, Medicine Evaluation Board and competent authorities of the concerned Member States.

This safety report consists of:

- a list of all suspected (unexpected or expected) serious adverse reactions, along with an aggregated summary table of all reported serious adverse reactions, ordered by organ system, per study;
- a report concerning the safety of the subjects, consisting of a complete safety analysis and an evaluation of the balance between the efficacy and the harmfulness of the medicine under investigation.

7.3 Follow-up of adverse events

All adverse events will be followed until they have abated, or until a stable situation has been reached. Depending on the event, follow up may require additional tests or medical procedures as indicated, and/or referral to the general physician or a medical specialist.

8. STATISTICAL ANALYSIS

8.1 Descriptive statistics

Descriptive statistics will be provided for the OQ scores per session (mean scores and standard deviations), the characteristics of the patient and therapist samples (proportions of males/female, mean, range and standard deviation of age, mean years of experience, proportion of patients with previous treatment) and patient diagnosis (proportion of patients with comorbidity, frequency of diagnostic categories).

8.2 Multivariate analysis

The data will be analysed using a three-level random intercept, random slope multilevel model, with repeated measures at level 1, patients within therapists at level 2 and therapists at level three. Randomisation takes place at the patient level (level 2). The OQ-45 total score is the dependent variable. The treatment conditions will be recoded into two dummy variables. Initial severity on the OQ-45 at start of treatment will be included in the model as a covariate for both the intercept and slope prediction.

8.3 Interim analysis

Interim power analysis will be performed after 6 months to test if an appropriate number of patients is included.

9. ETHICAL CONSIDERATIONS

9.1 Regulation statement

This study will be performed according to the Ethical Principles of Psychologists and Code Of Conduct by the American Psychological Association (2002).

9.2 Recruitment and consent

Patients are informed about the study at the end of the intake by the intake therapist or research assistant. Patients that are eligible for individuale treatment will receive the study information through mail and are contacted a week later by one of the research assistants or the researcher. On the phone they are asked if they have questions about the study and if they are willing to participate. If they are willing to participate the diagnostic interview is scheduled. This is usually planned one or two weeks in advance. At the beginning of the diagnostic interview, the consent form is signed by tht patient.

9.3 Benefits and risks assessment

The purpose of this study is to test a feedback system for therapists, with which they will be able to identify patients that are at risk for treatment failure earlier in treatment. From the studies of Michael Lambert, it has been shown that outcomes are greatly improved for the at risk patients when the feedback is used. So the potential benefit is that at risk patients (approximately 25%) are identified earlier and might receive better care. The risks of participating in the study are minimal. Patients are interviewed once and complete questionnaires, with a maximum of 17 times. The maximum time investment of the patient is 5 hours in total. Patients might experience unpleasant feelings while completing the questionnaires or while participating in the interview. These feelings can be addressed in treatment. If a patient experiences a great amount of distress during the diagnostic interview and needs to see a therapist directly, the psychiatric nurse from the crisis team that is on call at that moment, can always be asked to see the patient directly and the psychiatrist that is on call can be consulted as well.

10. ADMINISTRATIVE ASPECTS AND PUBLICATION

10.1 Handling and storage of data and documents

Most of the data is collected through the web-based software RequestXL, which is hosted online through a secure server. For the study, a specific module has been build, that is not available in the current program. Therapists, patients, researchers and research assistants are given access to relevant parts of the program by the principal investigator:

- Patients can only log into the program for completing the OQ-45 and ASC. After completing the questionnaire, they are logged out automatically. They cannot access any data or scores.
- Therapists have access to their own patients, but not those of other therapists. They can log in to view the feedback (if the patient is in one of the two feedback conditions) and to complete their questionnaires.
- Research assistants have access to all patients in their own research location. They can add new patients, erase patients and adapt patient information (not questionnaires).
- The principal investigator has access to all patients in the study, at all research locations and is the only person that can export the data. After study termination the data stored in Request XL will be cleared and deleted.

Data is exported to an Excel file and will be imported into SPSS. The data file contains patient number (which offers the opportunity to match the data with patient file data (e.g. diagnosis), but patient name, date of birth and other personal details that may identify the patient are removed. After matching the data the patient number will be removed. Data is stored on the server of GGZ Noord-Holland-Noord. After the study, the datafile is kept for 20 years, for the purpose of future data analysis.

The MINIPlus is administered through QuestManager, the questionnaire system that is used by the collaborating mental health care institutions in the SynQuest collaboration (GGZ Noord-Holland-Noord, Dimence, Rivierduinen, Reinier van Arkel groep, GGZ Breburg, GGZ Delfland, LUMC). Data from this system will not be used directly, but rather the diagnosis from the MINI is copied into the Electronisch Patient File.

The SCID-II PQ is scored on paper. The scoring sheet is kept in a locked closet and only contains the patient number as an identifier for the patient.

10.2 Amendments

Amendments are changes made to the research after a favourable opinion by the accredited METC has been given. All substantial amendments will be notified to the METC and to the competent authority. Non-substantial amendments will not be notified to the accredited METC and the competent authority, but will be recorded and filed by the sponsor.

10.3 Annual progress report

The sponsor/investigator will submit a summary of the progress of the trial to the accredited METC once a year. Information will be provided on the date of inclusion of the first subject, numbers of subjects included and numbers of subjects that have completed the trial, serious adverse events/ serious adverse reactions, other problems, and amendments.

10.4 End of study report

The investigator will notify the accredited METC of the end of the study within a period of 8 weeks. The end of the study is defined as the last patient's last six-months follow-up measure with the OQ-45. In case the study is ended prematurely, the investigator will notify the accredited METC, including the reasons for the premature termination.

Within one year after the end of the study, the investigator/sponsor will submit a final study report with the results of the study, including any publications/abstracts of the study, to the accredited METC.

10.5 Public disclosure and publication policy

Results will be published in an international peer reviewed journal, as well as in a national journal. In addition, in both participating mental health care organizations the final report will be distributed and results will be presented to the participating therapists. Patients will be informed through the patient newsletter, for which they can subscribe.

11. REFERENCES

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, NY: W.H. Freeman & Company.

Berking, M., Orth, U. & Lutz, W. (2006). Effects of systematic feedback to the therapist on patient progress. An empirical study in a cognitive-behavioral inpatient setting (in German). *Zeitschrift für Klinische Psychologie und Psychotherapie, 35,* 21-9.

Clairborn, C. D., & Goodyear, R. K. (2005). Feedback in psychotherapy. *Journal of clinical psychology: In Session, 61*(2), 209/217.

First, M.B., Spitzer, R.L., Gibbon, M., Williams, J.B.W. & Benjamin, L. (1997) *Structured Clinical Interview for DSM-IV Axis II Disorders (SCID-II).* Arlington, VA: American Publishing Inc.

Hannan, C., Lambert, M. J., Harmon, C., Nielsen, S. L., Smart, D. W., Shimokawa, K., et al. (2005). A lab test and algorithms for identifying clients at risk for treatment failure. *Journal of Clinical Psychology/ In Session, 61*(2), 155-163.

Hansen, N. B., Lambert, M. J., & Forman, E. M. (2002). the psychotherapy dose-response effect and its implications for treatment delivery services. *Clinical psychoogy: science and practrice, 9*(3), 329-343.

Harmon, S.C., Lambert, M.J., Smart, D.A., Hawkins, E., 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.

Hattie, J. & Timperley, H. (2007). The power of feedback. *Review of educational research, 77*, 81-112.

Hawkins, E. J., Lambert, M. J., Vermeersch, D. A., Slade, K. L., & Tuttle, K. C. (2004). The therapeutic effects of providing patient progress information to therapists and patients. *Psychotherapy research, 14*(3), 308-327.

Herold, D.M., Parsons, C.K. & Rensvold, R.V. (1996). Individual differences in the generation and processing of performance feedback. *Educational and psychological measurement, 56*, 5-25.

Herold, D.M. & Fedor, D.B. (1998). Individual's interaction with their feedback environment: The role of domain-specific individual differences. In G. Ferris (Ed.), *Research in personnel and human resources management, vol. 6* (pp. 215-254).

Herold, D. M., & Fedor, D. B. (2003). Individual differences in feedback propensities and training performance. *Human resource management review, 13*, 675-689.

Herold, D.M., Parsons, C.K. & Rensvold, R.V. (1996). Individual differences in the generation and processing of performance feedback. *Educational and psychological measurement, 56*, 5-25.

Hollenbeck, J. R. and H. J. Klein (1987). Goal commitment and the goal-setting process: problems, prospects, and proposals for future research. *Journal of Applied Psychology* **72**(2): 212-220.

Jamtvedt, G., Young, JM, Kristoffersen, DT, O'Brien, MA & Oxman, AD (2006). Audit and feedback: effects on professional practive and health care outcomes (review). *Cochrane database of systematic revie*ws 2006, Issue 2. Art. No.: CD000259. DOI:10.1002/14651858.CD000259.pub2.

Jong, K. , Moerbeek, M., Leeden, R. van der (2009). A priori power analysis in longitudinal three-level multilevel models: An example with therapist effects.

Jong, K. de, Lambert, M.J., Nugter, M.A. & Burlingame, G. (2008). Handleiding voor de Nederlandse Outcome Questionnaire (OQ-45). Salt Lake City, UT: OQ Measures LLD

Kluger, A.N. & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, meta-analyis and preliminary feedback intervention theory. *Psychological bulletin, 119*, 254-284.

Kluger, A.N. & DeNisi, A. (1998). Feedback interventions: Toward the understanding of a double-edged sword. *Current directions in psychological science, 7*, 67-72.

Lambert, M. J. (2007). Presidential adress: What we have learned from a decade of research aimed at improving psychotherapy outcome in routine care. *Psychotherapy research, 17*(1), 1-14.

Lambert, M.J. & Ogles, B.M. (2004). The efficacy and effectiveness of psychotherapy. In Lambert, M.J. (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (5th edition, pp. 139-193). New York: Wiley.

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 patients at risk for treatment failure. *Clinical psychology and psychotherapy, 9*, 149-164.

Lambert, M. J., Whipple, J. L., Smart, D. W., Vermeersch, D. A., Nielsen, S. L., & Hawkins, E. J. (2001). The effects of providing therapists with feedback on patient progress during psychotherapy: Are outcomes enhanced? *Psychotherapy research, 11*(1), 49-68.

Lecrubier, Y., Sheehan, D. V., Weiller, E., Amorim, P., Bonora, I., Sheehan, K. H. et al. (1997). The Mini International Neuropsychiatric Interview (MINI). A short diagnostic structured interview: reliability and validity according to the CIDI. *European Psychiatry, 12,* 224-231.

Lueger, R. L., Howard, K. I., Martinovic, Z., Lutz, W., Anderson, E. E., & Grissom, G. (2001). Assessing Treatment progress of individual patients using expected treatment response models. *Journal of Consulting & Clinical Psychology, 69*(2), 150-158.

Lutz, W., Lambert, M. J., Harmon, C. S., Tschitsaz, A., Schürch, E., & Stulz, N. (2006). The probability of treatment, success, failure and duration- what can be learned from empirical data to support decision making in clinical practice? *Clinical psychology and psychotherapy, 13*, 223-232.

Lutz, W., Martinovich, Z., Howard, K.I. (1999). Patient profiling: An application of random coefficient regression models to depicting the response of a patient to outpatient psychotherapy. *Journal of Consulting and Clinical Psychology, 67*, 571-577.

Riemer, M. & Bickman, L. (2003). The contextualized feedback intervention theory: Using basic psychological research to improve mental health services. Vanderbilt University: Unpublished manuscript.

Sapyta, J., Riemer, M., & Bickman, L. (2005). Feedback to clinicians: theory, research and practice. *JCLP/ In Session, 61*(2), 145-153.

Sheehan, D. V., Janavs, J., Baker, R., Harnett-Sheehan, K., Knapp, E., Sheehan, M. et al. (1998). MINI - Mini International Neuropsychiatric Interview - English Version 5.0.0 - DSM-IV. *Journal of Clinical Psychiatry, 59,* 34-57.

Veloski, J., Boex, J.R., Grasberger, M.J., Evans, A. & Wolfson, D.B. (2006). Systematic review of the literature on assessment, feedback and physicians' clinical performance: BEME Guide No. 7. *Medical teacher, 28*, 117-128.

Vliet, I. M. van, Leroy, H., & van Megen, H. J. G. M. (2000). *M.I.N.I. International Neuropsychiatric Interview; Dutch version 5.0.0.-R*. Department of Psychiatry University Medical Center Utrecht, the Netherlands.

Weertman, A., Arntz, A. & Kerkhofs, M.L.M. (2000). *SCID II. Gestructureerd Klinisch Interview DSM-IV As-II Persoonlijkheidsstoornissen.* Amsterdam: Pearson Assessment and Information

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, 50*(1), 59-68.