

# **Assessing Patient Progress in Psychological Therapy Through Feedback in Supervision: the MeMOS\* Randomized Controlled Trial (\*Measuring and Monitoring clinical Outcomes in Supervision: MeMOS)**

Kate M. Davidson

*Institute of Health and Wellbeing, University of Glasgow*

Michelle L. Rankin

*Institute of Health and Wellbeing, University of Glasgow*

Amelie Begley

*Institute of Health and Wellbeing, University of Glasgow*

Suzanne Lloyd

*Robertson Centre for Biostatistics, Institute of Health and Wellbeing, University of Glasgow*

Sarah J.E. Barry

*Robertson Centre for Biostatistics, Institute of Health and Wellbeing, University of Glasgow*

Paula McSkimming

*Robertson Centre for Biostatistics, Institute of Health and Wellbeing, University of Glasgow*

Lisa Bell

*Institute of Health and Wellbeing, University of Glasgow*

---

Correspondence to Kate M. Davidson, Institute of Health and Wellbeing, University of Glasgow, Glasgow, UK.  
E-mail: [kate.davidson@glasgow.ac.uk](mailto:kate.davidson@glasgow.ac.uk)

© British Association for Behavioural and Cognitive Psychotherapies 2017

Carole Allan

*NHS Greater Glasgow and Clyde*

Morag Osborne

*NHS Greater Glasgow and Clyde*

George Ralston

*NHS Greater Glasgow and Clyde*

Geraldine Bienkowski

*NHS Education for Scotland*

John Mellor-Clark

*Centre for Community Mental Health, Birmingham City University*

Andrew Walker

*Robertson Centre for Biostatistics, Institute of Health and Wellbeing, University of Glasgow*

**Background:** Psychological therapy services are often required to demonstrate their effectiveness and are implementing systematic monitoring of patient progress. A system for measuring patient progress might usefully ‘inform supervision’ and help patients who are not progressing in therapy. **Aims:** To examine if continuous monitoring of patient progress through the supervision process was more effective in improving patient outcomes compared with giving feedback to therapists alone in routine NHS psychological therapy. **Method:** Using a stepped wedge randomized controlled design, continuous feedback on patient progress during therapy was given either to the therapist and supervisor to be discussed in clinical supervision (MeMOS condition) or only given to the therapist (S-Sup condition). If a patient failed to progress in the MeMOS condition, an alert was triggered and sent to both the therapist and supervisor. Outcome measures were completed at beginning of therapy, end of therapy and at 6-month follow-up and session-by-session ratings. **Results:** No differences in clinical outcomes of patients were found between MeMOS and S-Sup conditions. Patients in the MeMOS condition were rated as improving less, and more ill. They received fewer therapy sessions. **Conclusions:** Most patients failed to improve in therapy at some point. Patients’ recovery was not affected by feeding back outcomes into the supervision process. Therapists rated patients in the S-Sup condition as improving more

and being less ill than patients in MeMOS. Those patients in MeMOS had more complex problems.

*Keywords:* Clinical supervision, patient progress, psychotherapy outcome research, CORE, randomized controlled trial

## Introduction

Giving clinicians feedback on the progress of patients with mental health difficulties may improve the outcome of therapy. The way in which feedback is given may influence therapy and therefore clinical outcome for patients. In one review, continuous feedback of outcomes was found in only half of twelve studies and helped short, but not longer-term, outcomes (Knaup et al., 2009). Intermittent feedback was only found to be effective for patients with psychiatric disorders who were not improving in therapy (de Jong et al., 2012). Research on continuous monitoring and feedback of clinical outcomes in psychological therapy has shown improvement in clinical outcomes in counselling settings, particularly for clients at risk of treatment failure, when feedback is given to therapists during therapy (Lambert et al., 2001; Lambert et al., 2002; Whipple et al., 2003; Harmon et al., 2007; Slade et al., 2008; Shimokawa et al., 2010). A systematic review of continuous feedback (Davidson et al., 2015) found that most of the research was conducted in the context of the United States health care system, by a small team of researchers and with patients with a narrow range of mental health problems and severity of illness. Patients in previous studies were mostly women, aged in their early 20s, attending university counselling services with mild mood disorders. In addition, no information was given on whether clients maintain benefits in the long term. It is therefore difficult to generalize these findings to other health settings, particularly for patients with more severe problems and with a greater range of health disorders. Three recent studies recruited patients with more severe psychiatric disorders: substance abuse (Crits-Christoph et al., 2012), psychiatric out-patients (Simon et al., 2012) and in-patients with eating disorders (Simon et al., 2013). When more severely unwell samples are investigated, more clients are rated as not improving with therapy and a smaller effect size is noted for feedback compared with the counselling services patients (Simon et al., 2012, 2013).

When Clinical Support Tools (CST) are used for clients who are Not on Track (NOT), outcomes have improved and this improvement is greater than giving progress feedback alone (Whipple et al., 2003; Harmon et al., 2007; Slade et al., 2008; Crits-Christoph et al., 2012; Simon et al., 2012, 2013). However, later studies showed smaller effect sizes than earlier studies (Simon et al., 2012, 2013).

We investigated the effect of supervision on patient informed clinical outcomes. This is a new field of research and a recent qualitative paper suggested that Clinical Outcomes for Routine Evaluation (CORE) outcomes, a system for measuring patient progress, can 'inform supervision' (Macdonald and Mellor-Clark, 2014). Given that therapists may not be alert to treatment failure (Hannan et al., 2005), may lack the ability to reflect on their level of therapeutic skill (Kraus et al., 2011; Walfish et al., 2012), tend to have an overly optimistic view of how their patients are progressing in psychological therapy, even when provided with evidence to the contrary (Lambert and Shimokawa, 2011), clinical supervision may be a more optimal method through which to provide patient informed clinical outcomes than therapists making a judgement. Feedback intervention theory suggests that feeding back

results of patient outcomes to health care staff might lead them to alter their treatment or use alternative treatments to improve care (Carlier et al., 2012). Therapeutic assessment theory suggests that there may be therapeutic effects of feedback in itself (Carlier et al., 2012). It is also possible that feeding back clinical outcomes to therapists and patients may change the dynamic between them and increase involvement in care with the potential for increased benefit. Supervision is integral to professional practice: its goal is to help improve knowledge and skills and to get feedback by discussing clinical practice with another more senior member of the profession. Feeding back patient clinical outcomes to a supervisor may enhance therapist–patient management by promoting more effective or efficient clinical practice and therefore improve patient care.

The present study aimed to determine whether monitoring clinical outcomes and providing feedback on patient progress would improve clinical outcomes for patients with a wide range and severity of mental health difficulties who were receiving psychological therapies in a city-wide psychological therapies service.

More specifically, we examined the impact of providing continuous clinical outcomes across therapy in two conditions. The first condition was Standard Supervision (S-Sup), where feedback on patient progress was only given to the therapist on a monthly basis for each patient and was *not* given to the therapist's supervisor, nor were immediate alerts sent to the therapist if a patient failed to progress. The second condition was the patient informed clinical outcomes supervision (Measuring and Monitoring clinical Outcomes in Supervision: MeMOS), where feedback on patient progress was given to both the therapist and supervisor, alerts were sent to both the therapist and supervisor and it was expected that patient progress would be discussed during clinical supervision.

The main question investigated was: do patients improve more quickly in the MeMOS condition compared with the S-Sup condition? More specific questions were: (1) are patient clinical outcomes, including rates of Reliable Clinical Change (RCI), superior in MeMOS compared with Standard Supervision (S-Sup)?; (2) does the recovery rate (CORE-10) show greater improvement in MeMOS for the subgroup of patients who fail to progress and therefore have an alert in MeMOS, compared with the recovery rate for those patients in S-Sup who fail to progress and thus would have met the criteria for an alert?; (3) do therapists overestimate progress less when feedback is given to supervisors and therapists (MeMOS vs S-Sup) condition?

## Method

The study design was a single-centre, stepped wedge, cluster-randomized controlled trial using three wedges (time periods), each consisting of the S-Sup condition followed by the MeMOS condition, with varying numbers of months in each wedge. Stepped wedge designs are used where an intervention is rolled out sequentially to the trial participants (either as individuals or clusters of individuals) over a number of time periods. The order in which the different individuals receive the intervention is determined at random and, by the end of the random allocation, all individuals or groups will have received the intervention (Brown and Lilford, 2006). Our design gave 24 months of recruitment of patients into both S-Sup and MeMOS. Therapist/supervisor dyads were randomized into one of the three wedges using a permuted-block method, stratified by service type.

**Setting.** Routine adult mental health and general medical services in a large health area in Scotland (NHS Greater Glasgow and Clyde).

**Participants.** All new patients who gave written informed consent after acceptance for psychological therapy were recruited. Patients were recruited from Primary Care (general practice 18%), Secondary Care (community specialist 17%) and Physical Health (31%), with the remainder from Specialist Services such as addictions, sexual abuse and neuropsychology.

Each patient received a study information sheet at initial assessment and indicated if they wished to take part in the study on their next appointment. The decision to take part in the study did not affect their treatment.

**Therapist/Supervisor Dyads** were the unit of randomization. Formal power calculations using the standard parallel group cluster power calculator ([www.abdn.ac.uk/hsru](http://www.abdn.ac.uk/hsru)) estimated that the study required nine therapists recruiting approximately 470 patients in total. Due to the nature of the stepped wedge design where all dyads should be present throughout the length of the study, the researchers thought it prudent to maximize recruitment of dyads due to potential loss to the study through promotion, maternity leave, etc. In total, 19 dyads, clinical psychologists and mental health practitioners working with adults, consented to take part. These dyads were representative of all psychology staff. A questionnaire was sent to all staff requesting the following details: age, gender, years qualified, area of clinical work, grade of post and number of people the individual supervised. Out of a possible 133 staff, 118 questionnaires were returned (89% return rate). No differences were found between the study dyads and other staff.

As mentioned above, 19 dyads volunteered to take part in MeMOS in month one of the study. The therapists began recruiting patients in month four following training in the use of measures and study set-up. Unfortunately, a number of dyads withdrew over the course of the study and three withdrew before recruitment of patients began. An additional five dyads were recruited and began recruiting patients in month seven of the trial. After 12 months, 12 dyads were still active and by the end of the patient recruitment phase of the trial, eight dyads remained. Withdrawal was due to reasons associated with therapists and not supervisors.

#### *Standard supervision (S-Sup) and MeMOS feedback conditions*

In the S-Sup condition, feedback on patient progress was given to therapists via email on a monthly basis only. Alerts (when a patient failed to improve or worsened) were *not* sent to therapists and they were *not* required to discuss patient progress during clinical supervision. Failure to improve or worsening was defined as no change in CORE-10 scores or an increase in a patient's score on CORE-10 across two therapy sessions (see CORE-10 in 'Measures' section). If requested, patients in the S-Sup condition also received feedback on their progress on a monthly basis via letter. In the MeMOS condition, feedback on patient progress was given to both therapist and supervisor on a monthly basis via email. Alerts were sent to both therapist and supervisor when a patient failed to improve or worsened according to their scores on CORE-10. The alert was sent as soon as the researchers received the patient's CORE-10 data. Therapists *were required* to discuss patients who had alerts during the next clinical supervision meeting, referring to a Clinical Support Tool (CST). The CST was devised by a group of senior clinicians to assist discussions in supervision about why a patient may have failed to improve in the MeMOS condition. Five main areas highlighted were diagnosis and formulation, therapy and goals, process vs content, medication and other factors – areas we

thought represented some common reasons for lack of progress in therapy. Patients in MeMOS received feedback on their progress on a monthly basis via letter. A questionnaire was given to dyads after each MeMOS patient had been discharged from therapy asking how much time they had spent discussing MeMOS patients during clinical supervision, how often and how useful was the CST and the usefulness of alerts.

Therapists therefore had continuous feedback of their patient's outcomes in both conditions (as they had access to patients' scores at all times). The two conditions, S-Sup and MeMOS, involved different ways of feeding back patient outcomes through supervision. The researchers fed back outcomes to the therapist (but not supervisor) in summary form on a monthly basis in S-Sup; this was to ensure that the therapist gave the patient outcomes at least some attention. In MeMOS, the researchers gave patient feedback monthly to therapist and supervisor. In addition, both therapist and supervisor received feedback on patient progress after an alert was triggered. Therapists therefore received continuous feedback in S-Sup and MeMOS conditions. Supervisors received intermittent feedback in the MeMOS condition (at the end of a month for all patients seen by the supervisee and immediately if a patient had an alert).

**Measures** The CORE system (Clinical Outcomes in Routine Evaluation) (Mellor-Clark et al., 2000) was chosen as it is well known throughout the UK as a generic measure of psychological distress and has been validated in samples within NHS primary care (Mellor-Clark et al., 2001), and in secondary care (Barkham et al., 2005). The system comprises three tools, CORE-Outcome Measure (CORE-OM), CORE Therapy Assessment Form (CORE-A) and CORE End of Therapy form (CORE-End). The CORE-OM is a 34-item generic pan-diagnostic measure of psychological distress with good internal and test-retest reliability. The CORE-OM was completed by patients at the beginning of therapy, end of therapy and at 6-month follow-up (from start of therapy), regardless of current treatment status. At 6-month follow-up, a patient could be in therapy, discharged or disengaged from therapy. Session-by-session monitoring was captured using the shorter 10-item version of the CORE-OM, CORE-10 (Barkham et al., 2013). Each item is scored on a five-point scale with higher scores indicating greater psychological distress. A score less than 11 is classified as below the threshold defining a clinical population.

Therapists also completed the Therapy Assessment Form (CORE-A) at initial assessment to collect referral information, demographics and data on presenting problem(s) or diagnosis, severity and duration. Therapists completed the End of Therapy Form (CORE-End) on discharge. This gave further information on the quantity and type of therapy used and patient problems and severity.

Data was stored electronically on the CORE NET program (Curtis Jenkins and Mellor-Clark, 2007), configured by the developers to meet the requirements of the study. This allowed data to be uploaded by the research team or by therapists participating in the study and provided therapists online access to their patients' CORE scores for the previous month only. The research team had access to all data to allow regular monitoring of patient progress and analyses at the end of the study.

### *Clinical Global Impression Scale*

Therapists provided an estimate of overall severity of illness and rate of improvement after each therapy session using the Clinical Global Impression (CGI) scale (Guy, 1976). The CGI Scale has good inter-rater reliability (Lyons Reardon et al., 2002).

### *Procedure*

**S-Sup condition.** Following consent, the therapist completed CORE-A and the patient, CORE-OM. Thereafter, the patient completed a CORE-10 at each therapy session and the therapist completed the CGI. Therapists were asked to complete the CGI without knowledge of the patient's scores on CORE-10.

The research assistant collected this information from therapists weekly as paper forms were used by the majority of therapists. Data were entered into the CORE Net application, giving an electronic record of each patient's CORE questionnaire scores and a progress chart. This chart was printed each month and formed the basis of the feedback to therapists throughout the duration of the study.

**MeMOS condition.** The same procedure as S-Sup condition was followed except that feedback charts were also sent to supervisors monthly. In MeMOS, patient's CORE-10 scores were monitored by the research team and when a patient failed to improve or worsened over two consecutive therapy appointments an alert was sent by email to the patient's therapist and their supervisor, drawing attention of both therapist and supervisor to the lack of improvement in the patient's scores.

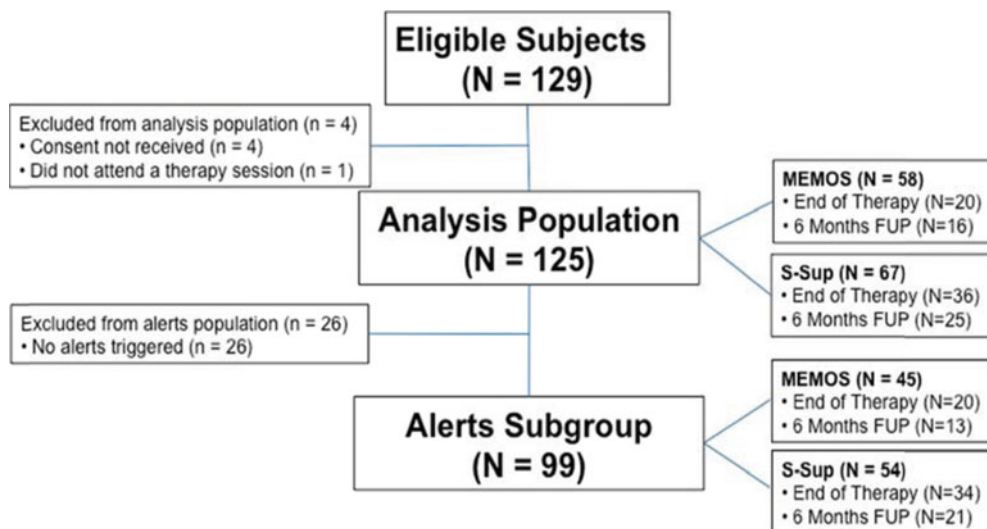
**Clinical Support Tool (only used in MeMOS).** Therapists and supervisors were asked to discuss possible reasons why the patient had not improved in therapy and to use the Clinical Support Tool (CST) to assist with this discussion (available from first author). The researchers collated information about the length of time dyads spent discussing each patient and the degree of usefulness of the alerts and the CST.

### *Statistical analysis*

Baseline characteristics are summarized overall and by condition (MeMOS or S-Sup) as mean (SD) for continuous variables and number (per cent) for categorical variables. Differences between the conditions were tested using two-sample *t*-tests and chi-squared tests for the continuous and categorical variables, respectively.

We used mixed effects repeated measures regression models to assess differences in outcomes between the two conditions. CORE outcomes were analysed as change from baseline scores (visit value – baseline value) and CGI as observed scores. The time effect was modelled as a categorical variable, with baseline scores and study group included as covariates; therapists and patients were included as random effects. Differences in the time trends between study groups were tested by including study group by time interaction terms. To test whether the changes over time were different between the two conditions, the *p*-value for inclusion of the interaction terms was extracted from these models. As none of these interactions was statistically significant, only the overall treatment effect with corresponding 95% confidence interval and *p*-value is reported (i.e. from the model excluding the therapy session by condition interaction term). Models were additionally adjusted for age, sex, living status, psychotropic medications, risk of self-harm and speciality of the therapist. Similar results were obtained from models without the additional adjustments noted above. Repeated measures logistic regression models were fitted as described above for the outcome of reliable clinical change. The mean CORE scores were plotted separately for each treatment group to visually display differences between the two conditions.





**Figure 1.** (Colour online) Flow of patients in study

## Results

One hundred and twenty-nine out of 180 eligible patients (72%) agreed to participate in this study (see Fig. 1). Results are based on 125 patients (67 in S-Sup, 58 in MeMOS). Data were removed for four patients as consent was not recorded. Follow-up was poor in both groups, although worse within the MeMOS group: 34% of patients under the MEMOS condition completed the end of therapy questionnaire, compared with 54% under the S-Sup condition. The 6-monthly follow-up was completed by 28% of patients under the MeMOS condition and 37% under the S-Sup condition.

Table 1 describes demographic and characteristics of the sample. The sample was mostly female and white (13% had no recorded ethnic origin), in their early 40s and just under half were not in employment. More patients lived alone in the MeMOS condition compared with S-Sup. Therapists could list more than one of the patients' problems. The main problems were depression and anxiety/stress. The majority (54%) of patients had been prescribed psychotropic medication. In addition, the majority of patients had previously attended secondary services (85%). Those participating had been referred from general practice (47%) and the remainder from the secondary healthcare system.

Some differences were found between the two conditions in terms of the main problem declared by patients (addictions, eating disorder, self-esteem, physical and relationship problems), in psychotropic medication intake and in therapist's estimate of risk of self harm.

For those patients who completed therapy, the duration of therapy was similar under the two conditions (163 days under S-Sup and 155 days under MeMOS). MeMOS patients, however, had fewer sessions offered to them (9 vs 12) and attendance at these sessions was lower (88 vs 94%).



**Table 1.** Baseline characteristics of patients

Mean (SD)	All ( <i>n</i> = 125)	MeMOS ( <i>n</i> = 58)	S-Sup ( <i>n</i> = 67)	<i>p</i> -value for difference <sup>1</sup>
Age (years)	42.37 (13.2)	41.63 (10.7)	43.00 (15.1)	0.56
Sex				
Female <i>n</i> (%)	77 (61.6%)	39 (67.2%)	38 (56.7%)	0.23
Ethnicity <i>n</i> (%)				
White	107 (85.6%)	47 (81.0%)	60 (89.5%)	0.34
Other	2 (1.6%)	2 (3.4%)	0 (0.0%)	
Unknown	16 (12.8%)	9 (15.5%)	7 (10.4%)	
Employment status				
Unemployed <i>n</i> (%)	34 (40.5%)	22 (45.8%)	12 (33.3%)	0.25
Living arrangement* <i>n</i> (%)				
Lives alone	37 (31.9%)	23 (43.4%)	14 (22.2%)	0.04
Lives with partner/other family	75 (64.7%)	29 (54.7%)	46 (73.0%)	
Other accommodation	4 (3.4%)	1 (1.9%)	3 (4.8%)	
Main problem <i>n</i> (%)				
Anxiety/stress	79 (63.2%)	41 (70.7%)	38 (56.7%)	0.11
Depression	78 (62.4)	41 (70.7%)	37 (55.2%)	0.08
Eating disorders	33 (26.4%)	22 (37.9%)	11 (16.4%)	0.01
Psychosis	1 (0.8%)	1 (1.7%)	0 (0.0%)	0.28
Addictions	21 (16.8%)	15 (25.9%)	6 (8.9%)	0.01
Personality problems	7 (5.6%)	1 (1.7%)	6 (8.9%)	0.08
Trauma/abuse	27 (21.6%)	15 (25.9%)	12 (17.9%)	0.28
Bereavement	21 (16.8%)	12 (20.7%)	9 (13.4%)	0.28
Self-esteem	55 (44.0%)	35 (60.3%)	20 (29.8%)	0.00
Physical problems	33 (26.4%)	22 (37.9%)	11 (16.4%)	0.00
Relationship problems	42 (33.6%)	27 (46.5)	15 (22.4%)	0.00
Other	33 (26.4%)	24 (41.4%)	9 (13.4%)	0.06
Psychotropic medication				
Yes <i>n</i> (%)	67 (53.6%)	37 (63.7%)	30 (44.8%)	0.01
Therapist assessed risk				
Suicide	43 (37.1%)	16 (30.2%)	27 (42.9%)	0.16
Self harm	28 (24.4%)	7 (13.2%)	21 (33.9%)	0.01
Harm to others	11 (9.6%)	3 (5.7%)	8 (12.9%)	0.19
Baseline scores				
CORE-10 score	19.8 (7.8)	20.8 (7.3)	18.9 (8.2)	0.45
CORE-34 score	18.7 (7.2)	19.2 (6.8)	18.2 (7.6)	0.18

Results are presented as means (standard deviation) or number (per cent). \*More than one option could be chosen for living arrangement and identified problems. <sup>1</sup>*p*-value estimated from *t*-test for continuous variables and chi-square test for categorical variables.

#### *Rate of change in patients' progress in therapy: MeMOS versus S-Sup*

Table 2 gives the results for the CORE outcomes (adjusted for baseline differences) for all patients in the study and results for patients with alerts. No overall difference was found in the rate of change between the two conditions for any of these outcomes (i.e. the rate of

**Table 2.** Results for CORE and CGI for (1) all patients and (2) for patients with an alert comparing MeMOS to S-Sup (adjusted for age, sex, living status, psychotropic medication, risk of self-harm and speciality of therapist)

Outcome	<i>p</i> -value for interaction*	Treatment effect (95% confidence interval)	<i>p</i> -value
(1) Results for all patients			
CORE-OM change from baseline	0.7438	2.96 (−1.33, 7.26)	0.1716
CORE-10 change from baseline	0.7646	2.78 (0.63, 4.94)	0.0119
CGI Improvement Score	0.7209	1.12 (0.57, 1.67)	0.0001
CGI Illness Score	0.4542	1.15 (0.55, 1.74)	0.0003
(2) Results for patients with alerts			
CORE-OM change from baseline	0.9222	2.59 (−1.93, 7.12)	0.2544
CORE-10 change from baseline	0.9215	2.77 (0.55, 5.00)	0.0154
CGI improvement score	0.7284	1.15 (0.61, 1.68)	<0.0001
CGI illness score	0.4358	1.21 (0.62, 1.80)	0.0001

\*Test for differences in time effects between treatment groups.

progress was similar across the two conditions). [Figures 2](#) (CORE-OM) and [3](#) (CORE-10 at each therapy session) give a graphic description of change including 6-month follow-up. For the CORE-10 change from baseline scores, this effect was found to be significant, in favour of the S-Sup condition for all patients and for patients with alerts (or who would have had alerts).

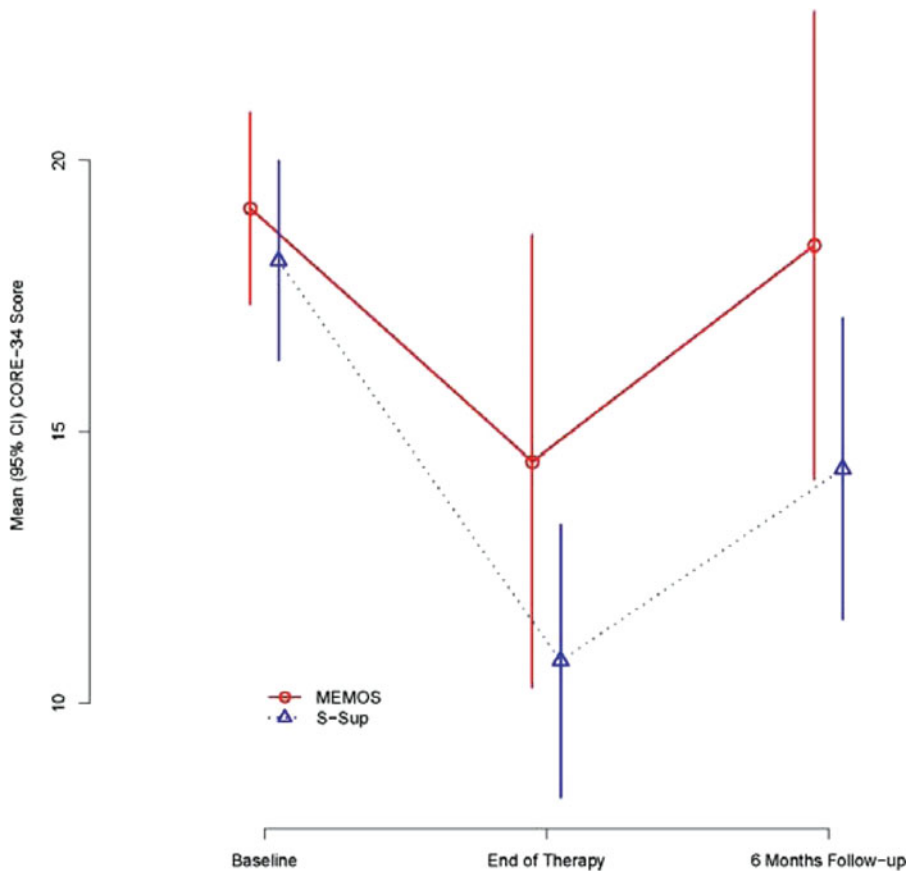
#### *Therapists' estimate of change over sessions and therapy*

There was no evidence of a difference in rate of change between the two conditions for either of the CGI outcomes. However, a significant condition effect was observed for both the CGI improvement and illness scores (see [Table 2](#)). Therapist ratings' of illness and improvement showed a perceived greater improvement and less illness in the S-Sup patients compared with MeMOS patients. Therapists considered that S-Sup patients had improved significantly more from sessions 1 to 3, and they remained significantly improved – but the patients did not improve any more compared with MeMOS patients subsequently.

#### *Reliable clinical change: MeMOS vs S-Sup*

The Reliable Change Index (RCI) was used to assess whether patients' clinical outcomes were superior in the MeMOS condition compared with the S-Sup condition. A decrease in clinical score on the CORE-OM of at least 5 points is considered a reliable clinical change (Barkham et al., 2006).

For CORE-OM at end of therapy, there were 23/35 (65.7%) patients in the S-Sup condition and 10/20 (50.0%) patients in the MeMOS condition that met the RCI criteria; there was no evidence of difference between the two conditions [OR: 0.52 (0.17, 1.60),  $p = 0.26$ ]. For CORE-OM at 6-month follow-up there were 11/23 (47.8%) patients in the S-Sup condition



**Figure 2.** (Colour online) Mean CORE-OM scores at baseline, end of therapy and at 6-month follow-up

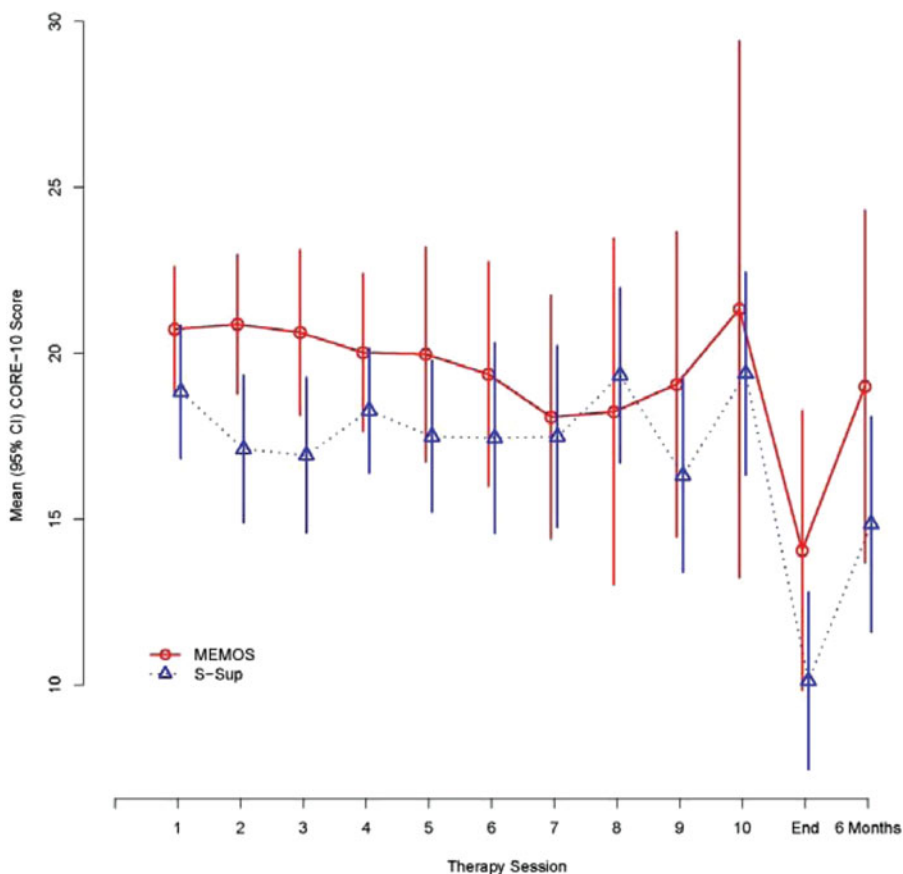
and 4/16 (25.0%) patients in the MeMOS condition that met the RCI criteria. There was no difference between the two conditions [OR: 0.36 (0.09, 1.47),  $p = 0.16$ ].

#### *MeMOS vs S-Sup for the sub-group of patients who triggered alerts*

Most patients (80%) failed to improve at some point during therapy. These patients triggered an alert during the MeMOS condition of the study. For the alerts subgroup, there was no difference between MeMOS and S-Sup in the rate of progress during therapy. Change of scores from baseline (CORE-10) were significantly better for S-Sup patients.

#### *Number of sessions*

The average number of therapy sessions delivered to patients in MeMOS was significantly lower than in S-Sup (5.9 vs 9.4:  $p < 0.001$ ). The MeMOS condition involved extra time with



**Figure 3.** (Colour online) Mean CORE-10 scores at each therapy session and at 6-month follow-up.

Due to the small number of patients having more than 10 therapy sessions, these plots only contain data for the first 10 therapy sessions, the final therapy session and the 6-month follow-up.

the supervisor (data for 47 patients in the MeMOS condition available: mean (SD) time spent discussing each MeMOS patient with an alert was 10.19 (15.46) (range 0 to 60 min). In the S-Sup condition, therapists confirmed that they did not discuss patients' outcomes feedback with their supervisor.

#### *Clinical supervision and use of CST*

Six out of eight dyads in MeMOS had patients with an alert. Of these, four therapists discussed patients during clinical supervision for an average of 10 min per patient. Three therapists referred to the CST during clinical supervision and all said they found the algorithm helpful.

## Discussion

This study investigated the impact of systematic monitoring and feedback of patients' progress in the context of supervision, in routine NHS psychological services across a broad range of health problems. Overall, patients progress in therapy and at 6 months follow-up did not differ between the two supervision conditions, standard supervision (S-Sup) and MeMOS. Nor was there evidence of a difference between the conditions in the number of patients who met reliable clinical change on the CORE-OM. There was a significant association between therapists' rating of improvement and patients' scores on session-by-session ratings (CORE-10) and therapists rated patients in the S-Sup condition as improving more and being less ill than patients in MeMOS.

The majority of patients failed to improve at some point during therapy. In the MeMOS condition, this triggered an alert. There was no difference in the overall progress between MeMOS and S-Sup for those patients who failed to progress in therapy from baseline to follow-up at 6 months. Again, therapists rated their patients as improving less and being more ill than patients in the S-Sup condition. Patients in the MeMOS condition were perceived as being more ill and improving less than those in S-Sup. Of note, more patients in the MeMOS group self harmed, had eating disorders and addictions. It may also suggest that when therapists and their supervisors receive alerts, therapists were more likely to be cautious about estimating patients' progress.

These finding differs from studies where patients had mild mental health problems but is similar to more recent studies and reviews of studies that have investigated patients with common mental health disorders (Kendrick et al., 2016) and more severe mental health problems (Davidson et al., 2015). Alerts have been previously aimed at 'not on track' patients with relatively mild problems and found to not be particularly helpful to those patients who are making progress (e.g. Lambert et al., 2001, 2002). We had a high identification rate of patients who failed to progress and it is possible that this number of alerts in too many cases decreased the effectiveness of the alerts, making them commonplace. We have added to this literature another layer of complexity by also feeding back patient results to supervisors. Our findings suggest that systematic monitoring of outcomes and their use in supervision is complex and outcome of therapy involves multiple patient, therapist and service factors that cannot be controlled for in a study in routine clinical practice.

Overall, patients attended for fewer sessions than would be recommended for effective treatment. This may have had an unknown impact on the overall outcome of the study. The number of therapy sessions was significantly lower for patients in the MeMOS condition. It is possible that feedback of patient progress to the supervisor and therapist affected decisions about the number of sessions offered. However, the drop-out rate was higher in this condition. It is possible that MeMOS patients had more complex problems and were more likely to give up therapy.

We faced some challenges conducting research in routine NHS care and implementing systematic monitoring of patient progress into routine care. Many patients disengaged from therapy early, making it difficult to collect follow-up data. More than half of the therapists left the study prematurely, which had a direct effect on the number of patients that were recruited. Changing jobs, maternity leave and illness were the main reasons. The stepped wedge design therefore presented some specific problems, and although anticipated, these

were more than expected. The under-recruitment means that the study was limited in its power to detect differences between the conditions.

It is also possible that our alert system may have been overly sensitive (alert triggered if a patient's CORE-10 score remained the same or worsened by one or more point on a session-to-session basis). Some therapists suggested that alerts should have been sent if a patient's score got worse than their baseline score. In addition, we asked therapists not to use CORE data in the S-Sup condition when discussing patients in supervision. Therapists stated they did not use this data. Nonetheless, they could have discussed patients in supervision and may have been more likely to discuss patients whom they know were failing to progress in therapy. All of the above factors may limit the outcome of the study and suggest caution in the interpretation of the findings; however, they also provide important information to researchers considering carrying out a future study into the impact of clinical supervision in routine psychological care. Future studies might wish to investigate what might be the optimal alert systems of clinical deterioration for therapists and supervisors, with a view to increasing therapeutic impact and efficiency of services.

There are strengths over previous studies – the trial took place in routine psychological services with patients from a wide range and severity of mental and physical health problems and the therapist/supervisor dyads were representative of a larger psychological service. Although there were more female patients than male, the age span was much greater than that of patients recruited in previous studies (ranging up to 76 years). Other strengths were the use of more than one outcome measure, patients could rate their own progress on a session-by-session basis and therapists rated the progress of their patients throughout the course of therapy. We also investigated the long-term benefits of monitoring patient progress by including a 6-month-follow-up.

### Acknowledgements

The authors thank the staff and patients of NHS Greater Glasgow and Clyde who took part in the study.

*Financial support:* This work was supported by the Chief Scientist's Office, Scottish Government Health Directorates, grant references: CZH/4/708 and CZH/4/804.

*Conflicts of interest:* None.

*Ethical standards:* Ethical approval for the study was obtained from the West of Scotland REC 5 committee 2.

### References

- Barkham, M., Bewick, B., Mullin, T., Gilbody, S., Connell, J., Cahill, J., Mellor-Clark, J., Richards, S., Unsworth, G. and Evans, C. (2013). The CORE-10: A short measure of psychological distress for routine use in the psychological therapies. *Counselling and Psychotherapy Research*, 13, 3–13.
- Barkham, M., Gilbert, N., Connell, J., Marshall, C. and Twigg, E. (2005). Suitability and utility of the CORE-OM and CORE-A for assessing severity of presenting problems in psychological therapy services based in primary and secondary care settings. *British Journal of Psychiatry*, 186, 239–246.

- Barkham, M., Mellor-Clark, J., Connell, J. and 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 and Psychotherapy Research*, 6, 3–15.
- Brown, C. and Lilford, R. J.** (2006) The stepped wedge trial design: a systematic review. *BMC Medical Research Methodology*, 6, 54. [www.biomedcentral.com/1471-2288/6/54/prepub](http://www.biomedcentral.com/1471-2288/6/54/prepub)
- Carrier, I. V. E., Meuldijk, D., Van Vliet, I. M., Van Fenema, E., Van der Wee, N. J. A. and Zitman, F. G.** (2012). Routine outcome monitoring and feedback on physical or mental health status: evidence and theory. *Journal of Evaluation in Clinical Practice*, 18, 104–110. doi:10.1111/j.1365-2753.2010.01543.x
- Crits-Christoph, P., Ring-Kurtz, S., Hamilton, J. L., Lambert, M. J., Gallop, R., McClure, B., Kulaga, A. and Rotrosen, J.** (2012). A preliminary study of the effects of individual patient-level feedback in outpatient substance abuse treatment programs. *Journal of Substance Abuse Treatment*, 42, 301–309.
- Curtis Jenkins, A. and Mellor-Clark, J.** (2007). CORE NET [computer program]. Available at: [http://www.coreims.co.uk/Buy\\_Net\\_Software.html](http://www.coreims.co.uk/Buy_Net_Software.html) (accessed 18 January 2015).
- Davidson, K., Perry, A. and Bell, L.** (2015). Would continuous feedback of patient's clinical outcomes to practitioners improve NHS psychological therapy services? Critical analysis and assessment of quality of existing studies. *Psychology and Psychotherapy: Theory, Research and Practice*, 88, 21–37. doi: 10.1111/papt.12032
- de Jong, K., van Sluis, P., Nugter, M. A., Heiser, W. J. and 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.
- Guy, W.** (1976) Clinical Global Impressions. In *ECDEU Assessment Manual for Psychopharmacology*, revised (DHEW Publ. No. ADM 76-338), pp. 218–222. National Institute of Mental Health: Rockville, MD.
- Hannan, C., Lambert, M. J., Harmon, C., Nielson, S. L., Smart, D. W., Shimokawa, K. and Sutton, S. W.** (2005). A lab test and algorithms for identifying clients at risk of treatment failure. *Journal of Clinical Psychology*, 61, 155–163.
- Harmon, S. C., Lambert, M. J., Smart, D. M., Hawkins, E., Nielsen, S. L., Slade, K. and Lutz, W.** (2007). Enhancing outcome for potential treatment failures: Therapist-client feedback and clinical support tools. *Psychotherapy Research*, 17, 379–392.
- Kendrick, T., El-Gohary, M., Stuart, B., Gilbody, S., Churchill, R., Aiken, L., Bhattacharya, A., Gimson, A., Brütt, A. L., de Jong, K. and Moore, M.** (2016). Routine use of patient reported outcome measures (PROMs) for improving treatment of common mental health disorders in adults. *Cochrane Database of Systematic Reviews*, 7, CD011119. doi: 10.1002/14651858.CD011119.pub2
- Knaup, C., Koesters, M., Schoefer, D., Becker, T. and Puschner, B.** (2009). Effect of feedback of treatment outcome in specialist mental healthcare: Meta-analysis. *British Journal of Psychiatry*, 195, 15–22.
- Kraus, D. R., Castonguay, L., Boswell, J. F., Nordberg, S. S. and Hayes, J. A.** (2011). Therapist effectiveness: implications for accountability and client care. *Psychotherapy Research*, 21, 267–276.
- Lambert, M. J. and Shimokawa, K.** (2011). Collecting client feedback. *Psychotherapy*, 48, 72–79.
- Lambert, M. J., Whipple, J. L., Smart, D. W., Vermeersch, D. A. and Nielsen, S. L.** (2001). The effects of providing therapists with feedback on patient progress during psychotherapy: Are outcomes enhanced? *Psychotherapy Research*, 11, 49–68.
- Lambert, M. J., Whipple, J. L., Bishop, M. J., Vermeersch, D. A., Gray, G. V. and 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.



- Lyons Reardon, M., Cukrowicz, K. C., Reeves, M. D. and Joiner, T. E.** (2002). Duration and regularity of therapy attendance as predictors of treatment outcome in an adult outpatient population. *Psychotherapy Research*, 12, 273–285.
- Macdonald, J. and Mellor-Clark, J.** (2014). Correcting psychotherapists' blindsidedness: formal feedback as a means of overcoming the natural limitations of therapists. *Clinical Psychology and Psychotherapy*, 22, 249–257. doi: [10.1002/cpp.1887](https://doi.org/10.1002/cpp.1887)
- Mellor-Clark, J., Barkham, M., Connell, J. and Evans, C.** (2000). Practice-based evidence and need for a standardised evaluation system: Informing the design of the CORE System. *European Journal of Psychotherapy, Counselling and Health*, 2, 357–374.
- Mellor-Clark, J., Connell, J., Barkham, M. and Cummins, P.** (2001). Counselling outcomes in primary health care: a CORE system data profile. *European Journal of Psychotherapy and Counselling*, 4, 65–86.
- Shimokawa, K., Lambert, M., and Smart, D.** (2010). Enhancing treatment outcome of patients at risk of treatment failure: Meta-analytic and mega-analytic review of psychotherapy quality assurance system. *Journal of Consulting and Clinical Psychology*, 78, 298–311.
- Simon, W., Lambert, M. J., Busath, G., Vazquez, A., Berkeljon, A., Hyer, K. et al.** (2013). Effects of providing patient progress feedback and clinical support tools to psychotherapists in an inpatient eating disorders treatment program: A randomized controlled study. *Psychotherapy Research*, 23, 287–300.
- Simon, W., Lambert, M. J., Harris, M. W., Busath, G. and Vazquez, A.** (2012). Providing patient progress information and clinical support tools to therapists: Effects on patients at risk of treatment failure. *Psychotherapy Research*, 22, 638–647.
- Slade, K., Lambert, M. J., Harmon, S. C., Smart, D. W. and Bailey, R.** (2008). Improving psychotherapy outcome: The use of immediate electronic feedback and revised clinical support tools. *Clinical Psychology and Psychotherapy*, 15, 287–303.
- Walfish, S., McAlister, B., O'Donnell, P. and Lambert, M. J.** (2012). An investigation of self-assessment bias in mental health providers. *Psychological Reports*, 110, 639–644.
- Whipple, J. L., Lambert, M. J., Vermeersch, D. A., Smart, D. W., Nielsen, S. L. and Hawkins, E. J.** (2003). Improving the effects of psychotherapy: The use of early identification of treatment and problem-solving strategies in routine practice. *Journal of Counselling Psychology*, 50, 59–68.