An examination of the relationship between working alliance and outcome in Cognitive Behaviour Therapy and Counselling for Depression

Elizabeth Gilley

Submitted in partial fulfilment for the award of Doctor of Clinical Psychology at the University of Sheffield

November 2017

This page is left intentionally blank

Declaration

I declare that this work has not been submitted for any other degree at the University of Sheffield or any other institution. This thesis is my own original work and all other sources have been referenced accordingly.

This page is left intentionally blank

Word Count

Literature Review

Excluding references	7989
Including references	9868
Including references and appendices	11494

Research Report

Excluding references	9735
Including references	11830
Including references and appendices	16567

Total

Including references and appendices	28513

This page is left intentionally blank

Abstract

Literature Review

This literature review critically evaluates and synthesises the available literature on the relationship between working alliance (WA) and outcome of psychotherapy in Cognitive Behaviour Therapy (CBT) for depression and/or anxiety. An electronic database search was undertaken using terms related to WA, cognitive behaviour therapy, and outcome or symptom reduction. The final review included twenty studies. A narrative synthesis of the literature was undertaken. The strength of the alliance-outcome relationship varied across studies, with effect sizes ranging from small to large. In general, there was support for the presence of the alliance-outcome relationship at early, mid and late therapy. There was limited evidence that the alliance-outcome relationship may be more relevant for some therapies than others and that certain aspects of the therapy relationship may be of importance for CBT. The methodological limitations of the literature and recommendations for clinical practice and future research are discussed.

Research Report

This study investigates the WA in CBT and Counselling for Depression (CfD). Using a subset of participants from a wider trial (n = 40), the strength of the WA was rated for sessions one, three and five of psychotherapy by trained observers using the Working Alliance Inventory-Observer Form (WAI-O). The resulting scores were compared for the two therapies and their relationship to outcome (in addition to demographic and clinical variables) was examined. An investigation of the experience of coding the WA was also undertaken. Overall WAI scores and the subscales of Goal and Task were higher for CBT than CfD, with medium effect sizes, though not all results achieved statistical significance. For the full sample and CfD subsample, there was a lack of relationships between WA and the additional variables to outcome. For CBT, number of

sessions completed, WA, first session depression score explained 52% of the variance in final session depression scores. The coder questionnaires highlighted some positive aspects of the WA measure, but highlighted several issues with coding that could confound alliance scores. The clinical implications of these findings are discussed, and future research recommendations are made.

Acknowledgments

I would like to thank my supervisor, Professor Gillian Hardy for her help and support through this process. Also thanks to (soon to be Dr) Dave Saxon for his advice throughout the project. Thanks to Kerry Arderm for going above and beyond in getting the outcome data for me. A huge thanks to my coders, Eilish Pearson, Clinton Felicio, Raluca Iancu and Melanie Simmonds-Buckley: your enthusiasm for the project kept me going.

I would like to thanks my partner Matt for putting up with me throughout this process. Thanks also to Jo, who read through some of my drafts and gave me invaluable feedback. Finally, this is dedicated to my children: my daughter, Hazel, to baby bean and to the memory of Maxwell. I love you all.

This page is left intentionally blank

Table of Contents

Part Two: Research Report

Working alliance in Cognitive Behaviour Therapy and Counselling for Depression: A comparison of therapies and relationship to outcome

Research Report Appendices

Research Report Appendices Continued

This page is left intentionally blank

Part One: Literature Review

Examining the relationship between working alliance and outcome in Cognitive Behaviour Therapy for anxiety and depression

This page is left intentionally blank

Abstract

Objective: No review to date has examined the potential relationship between working alliance and outcome in Cognitive Behaviour Therapy (CBT) for anxiety and depression. This systematic literature review aimed to: 1) summarise the research evidence examining the alliance-outcome correlation in CBT for anxiety and depression; 2) examine the temporal relationship between alliance and symptom change in CBT for anxiety and depression; 3) investigate which aspects of the alliance are important in CBT for anxiety and depression.

Method: A systematic literature review was undertaken. Three databases were searched (PsychInfo, Web of Science Core Collection and Medline) on 01 March 2017. Search terms related to working alliance, cognitive behaviour therapy, and outcome or symptom reduction were used. The search was limited to English Language papers only but not limited by publication date.

Results: Twenty studies were included in the review. The strength of the alliance-outcome relationship varied across studies, with effect sizes ranging from small to large. There was limited evidence in favour of the reverse causality hypothesis. In general, however, there was support for the presence of the alliance-outcome relationship at early, mid and late therapy. There was limited evidence that the alliance-outcome relationship may be more relevant for some therapies than others and that certain aspects of the therapy relationship may be of particular importance for CBT.

Conclusions: Alliance may be important for the outcome of CBT for anxiety and depression, particularly the contributions of the client. Further research examining how client characteristics influence alliance and comparing alliance across therapies are required to understand the alliance-outcome relationship for CBT.

Practitioner Points

- CBT therapists working with clients with anxiety and depression would benefit from routinely using standardised measures of alliance with clients.

- CBT therapists working with clients with anxiety and depression would benefit from incorporating interventions to improve alliance as part of treatment.
- It might be of use for therapists to consider how aspects of a client's history, attachment style and interpersonal style might impact the alliance and to be aware that therapist's perception of the alliance may differ to that of the client.

Limitations

- Only research papers published in English were included, and there was no citation search or search of grey literature; all of which may bias the findings of the review.
- Despite attempts to narrow down the literature to specific interventions for specific presenting problems, there was still heterogeneity in the data included in the review in terms of presenting problems, intervention and treatment setting.

Depression and anxiety are amongst the most common mental health problems experienced in the United Kingdom (UK; NICE, 2011a). It is estimated that between four and ten percent of the population will experience major depression at some point in their lifetime, and over five percent will experience generalised anxiety disorder (NICE, 2011b). The comorbidity between Major Depressive Disorder (MDD) and anxiety disorders is high, with mixed anxiety and depression the cause of one fifth of working days lost in Britain (Das Munshi et al., 2008), adding an economic imperative to finding and understanding effective treatments.

The evidence base for the treatment of anxiety and depression is dominated by Cognitive Behaviour Therapy (CBT; e.g. Cuijpers et al., 2014). Current recommendations for treating anxiety and depression in the UK include CBT as a first line treatment (NICE, 2009; 2011b). CBT is based on the cognitive model of Beck (1970) which proposes that psychological distress is caused by maladaptive thoughts

which underlie problematic feelings and behaviours. The therapy consists of strategies, including challenging thoughts, problem solving, and use of behavioural techniques such as exposure (Gaudiano, 2008). The dominance of CBT is not without controversy, not least from within the CBT community itself (see Gaudiano, 2008). Consequently, "third wave" therapies have emerged (Hayes, 2004). Though based on CBT, third wave therapies diverge towards contextual explanations of difficulties (as opposed to mechanistic) and experiential (as opposed to taught) approaches to therapy (Hayes, 2004).

Despite the extensive research literature, on CBT there remains uncertainty as to the mechanisms by which CBT leads to change. Longmore and Worrell (2007) reviewed thirteen component studies and found no differences in outcomes between groups of participants assigned different CBT techniques. However, the authors did not combine the findings of the studies. A meta-analysis of 49 studies failed to find a significant relationship between therapist adherence to CBT techniques or therapist competence and treatment outcome (Webb, DeRubeis & Barber, 2010). This problem is not limited to CBT, with cross therapy meta-analyses finding limited evidence for the addition or removal of therapy components (Bell, Marcus & Goodland, 2013).

The lack of evidence for the "specific factors" that produce change in addition to the claim that all psychotherapies are equally effective (e.g. Wampold, Imel & Miller, 2009), has led to the search for the "common factors" that make psychotherapy successful. Though this approach has its detractors (e.g. Carroll & Roundsaville, 2010), several common factors have been identified (Tracey, 2010; Wampold, 2015). Of these, the Working Alliance (WA) is one of the most frequently examined areas in process research (Doran, 2014).

The WA is a component of the therapeutic relationship in psychotherapy (Castonguay, Constantino & Holtforth, 2006). Whilst ways of conceptualising the WA

have differed (Elvins & Green, 2008), it has generally been described as representing the collaborative parts of the therapeutic relationship, involving both therapist and client (Constantino, Castonguay & Schut, 2002). Bordin (1979) developed a pantheoretical definition of WA consisting of the therapist and client agreeing upon therapy goals (goal), negotiating how to achieve those goals (task), as well as the relationship between the therapist and client (bond). A number of pantheoretical measures have been developed in order to measure the WA for the purposes of research: completed by client, therapist or independent observers (Elvins & Green, 2008).

Cross therapy meta-analyses suggest that WA is positively associated with treatment outcome with correlation effect sizes ranging from .22 to .29 (Flückiger, Del Re, Wampold, Symonds & Horvath, 2012; Horvath & Symonds, 1991; Martin, Garske & Davis, 2000). Despite claims that the alliance-outcome correlation is ubiquitous (Flückiger, et al., 2012), there remains elements of the relationship that are not fully understood. Firstly, there is doubt as to the temporal nature of alliance, and its progression during therapy, with some researchers claiming that alliance is a product of symptom change, rather than vice versa (DeRubeis & Feeley, 1990). Though this finding has not always been replicated (e.g. Falkenström, Ekeblad & Holmqvist, 2016), it remains an area of uncertainty. While the most recent meta-analysis looked at the alliance outcome relationship at different stages of therapy (Flückiger et al., 2012), the authors did not examine the reverse causality hypothesis.

A second controversy in the alliance-outcome debate is its relevance in CBT (Flückiger et al., 2012). This is because while Beck recognised the importance of the therapeutic relationship in cognitive therapies, he did not regard it as sufficient to produce change in itself (Beck, Shaw, Rush & Emery, 1979). This has led some authors to claim that the relationship is fairly unimportant for CBT (Siev, Huppert & Chambless, 2009). This contrasts with other therapies from the Rogerian tradition,

which view the WA as a vehicle for change (Gelso & Hayes, 1998). It is therefore important to investigate whether the alliance-outcome relationship is present for CBT. In addition, it has been suggested that some aspects of alliance are more relevant to CBT than other therapies (Webb, DeRubeis, Amsterdam, Shelton & Hollon, 2011). Therefore, an investigation of whether specific aspects of alliance (as measured in some subscales of measures) are notably important in CBT would be beneficial.

Previous reviews and subsequent meta-analyses have reported significant heterogeneity in the data (e.g. Flückiger et al., 2012). It has been argued that overly blunt aggregation of data obscures important treatment effects (Siev, Huppert & Chambless, 2009). For example, there is some evidence that WA may be more important for some presenting problems than others (e.g. Barber et al., 1999). There is a need therefore to examine the alliance-outcome relationship for a narrower range of specific presenting problems and treatments.

Aims. To systematically review previous research investigating the relationship between WA and treatment outcome for CBT in people experiencing anxiety and depression. Specific objectives are to locate, appraise and where possible synthesise existing research examining:

1) the alliance-outcome correlation in CBT for anxiety and depression; 2) the temporal relationship between alliance and symptom change in the in CBT for anxiety and depression and 3) which aspects of the WA might be important in affecting outcomes in those receiving CBT for anxiety and depression.

Method

To address the aims and objectives, a systematic literature review was undertaken. This method was chosen because it is effective for locating, appraising and synthesising evidence in health care. Systematic literature reviews use pre-defined methods to

identify, evaluate and summarise relevant research data (Centre for Reviews and Dissemination, 2008). The process follows the scientific process and as such aims to use transparent methods and minimise bias (Moher et al., 2015). This review was conducted using the 12 step guidelines published by Kable, Pich and Maslin-Prothero (2012).

Study criteria for inclusion

Population. The included study population were adults experiencing anxiety or depression as outlined in the Diagnostic and Statistical Manual of Mental Disorders-5 (DSM-5: American Psychiatric Association, 2013). However, a formal diagnosis was not required for inclusion.

Intervention. The intervention was CBT or therapies within the CBT tradition, including third wave therapies.

Outcome. Included studies needed to report a standardised assessment of treatment outcome including measures of symptomology or outcome (e.g. quality of life measures).

Alliance. Alliance was measured using any standardised measure of working alliance.

Search Strategy

Three databases were searched (PsychInfo, Web of Science Core Collection and Medline) on 01 March 2017. Search terms for WA were based on those used by Horvath & Bedi (2002). The search terms used were: (helping alliance OR therap* alliance OR working alliance OR therap* relationship) AND (cognitive behavio* therapy OR CBT) AND (outcome OR symptom reduction). The search was limited to English Language papers only but not limited by publication date. A combination of subject headings and free text terms were used where available. The full search strategy is in Appendix A.

Screening and Selection

The initial search yielded 1069 papers. After duplicates were removed, paper titles and abstracts were assessed for relevance by the author. The remaining papers were subject to a full text review for eligibility. The process is outlined in the PRISMA diagram (Moher, Liberati, Tetzlaff, & Altman, 2009) in Figure 1.

Eligibility Criteria

The inclusion criteria were i) primary presenting problem of anxiety and/or depression disorders (ii) CBT intervention or variants within the cognitive behavioural tradition (see Table 1), (iii) outcomes assessed using a validated measure of outcome, (iv) WA assessed using a validated measure, (v) peer reviewed journal.

Exclusion criteria were as follows: (i) child participants (ii) inpatient settings, (iii) group therapy, (iv) therapy conducted via online, telephone or video-conference due to the possible differences these conditions may make to the WA (Horvath & Symonds, 1991; Preschl, Maercker & Wagner, 2011), (v) research with fewer than five participants due to the differences in underlying assumptions and design (Horvath & Symonds, 1991).

There were a number of papers that included comparator groups (e.g. group vs individual therapy; CBT vs IPT) that did not meet inclusion criteria for the review. Where this was the case, the paper was included if the paper reported findings separately for the group that met inclusion criteria. To reduce potential bias, where papers did not report findings separately, the lead author was contacted to request the data separately. There were no positive responses to this request.

Several of the papers reported analysis of the same data set. In these cases, the research paper that was most pertinent to the questions the review was addressing was included. Where this was unclear, the decision was made in conjunction with a second reviewer, who was also a third year Doctorate of Clinical Psychology student. In each case, both reviewers agreed upon the decisions made independently.

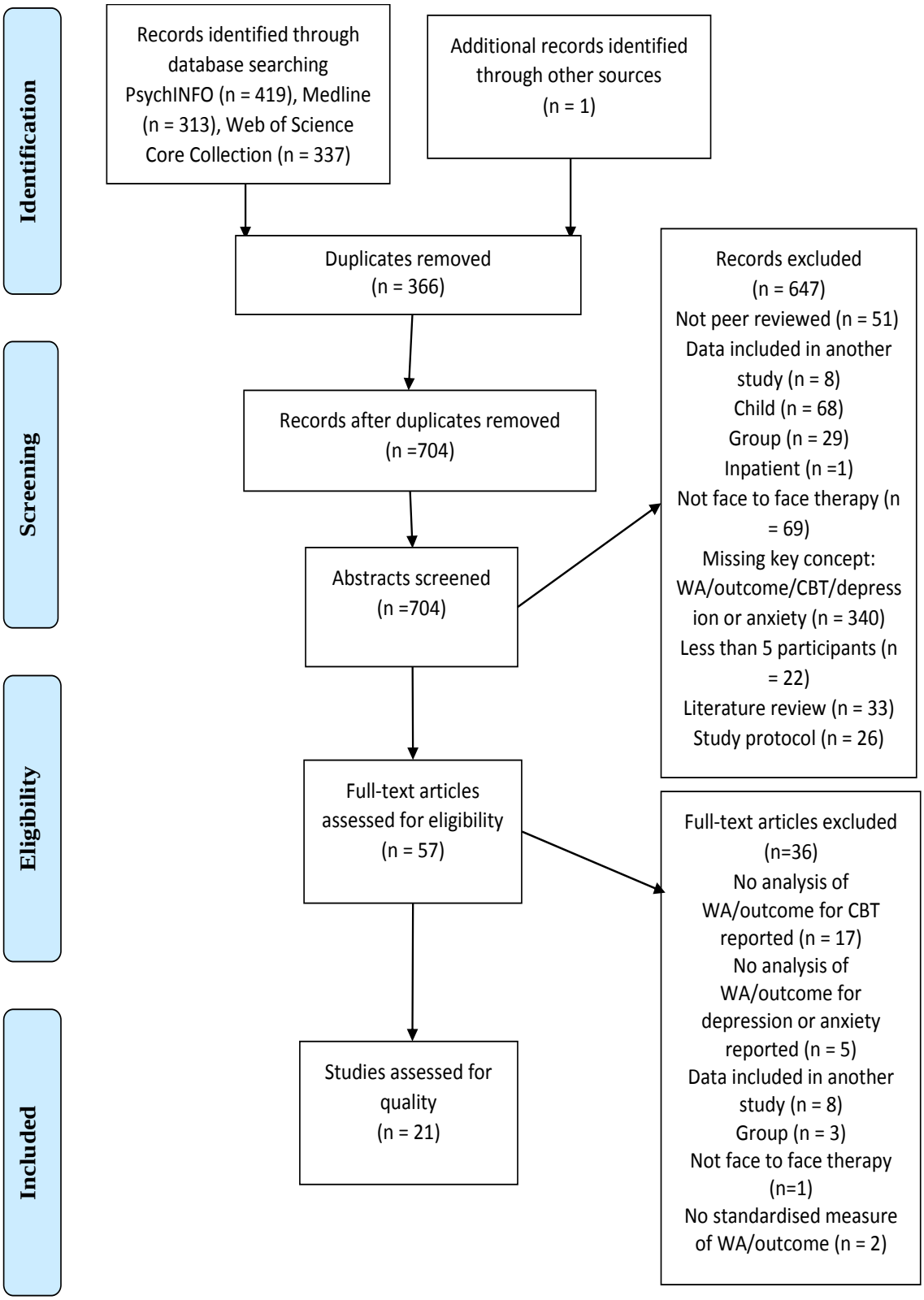

Figure 1. Literature search process

Quality Assessment and Data Synthesis

The quality of papers included in a literature review will impact on the reliability of its results (Centre for Reviews & Dissemination, 2008). Assessing the quality of papers can be used to weight the interpretation of findings where appropriate. Quality assessment was undertaken using the Downs and Black adapted checklist (Downs & Black, 1998; see Appendix B). This 27-item checklist is designed to evaluate the study quality of healthcare interventions. The checklist was selected because it can be used for both randomised and non-randomised studies and has high internal consistency, good test-retest and inter-rater reliability (Downs & Black, 1998). It assesses studies based on study quality, external and internal validity and statistical power. For each item, a score is assigned ("yes" = 1, "no" = 0, "unable to determine" = 0). Question 27 was adapted as follows: ("power calculation reported/sample sufficient" = 2, "power calculation reported/sample insufficient" = 1, "no power calculation reported" = 0). The final score is the sum of the items, with a maximum score of 28. A higher score indicates greater study quality. An arbitrary cut-off of ten was chosen as a minimum score for inclusion in the review.

Five of the papers were selected at random and subject to assessment by a second reviewer. The reviewer was a third year Doctorate of Clinical Psychology student who had experience of quality assessment using the Downs and Black (1998) checklist. A table of scores given by each rater is in Appendix D. Discrepancies in scoring between the first and second reviewer were discussed, and the first reviewers scores were ultimately accepted on each occasion. Inter-rater reliability was calculated using two-way mixed absolute, average measures intra-class correlation (ICC; MCGraw & Wong, 1996) to determine the extent to which the raters agreed. An excellent level of inter-rater reliability was found (ICC = 0.95; Cicchetti, 1994).

Data synthesis was undertaken in two stages. Firstly, the data was gathered

using data extraction forms. Effect sizes (r and R^2) were extracted from the papers (see table 2). There was insufficient data to convert R^2 to an r statistic. Effect sizes are reported according to Cohen (1988) where an r of .1 - .29 = small effect size, r of .3 - .49 = medium effect size and an r of $r \geq .5$ = large effect size.

The results were extracted: mean ratings of alliance, alliance in early therapy, alliance at mid therapy, alliance late in therapy, subscales of alliance measures and the reverse causality hypothesis. The definition of early, middle and late therapy is complicated by differences in research design. For the purposes of this review, early therapy is defined as sessions one to four, mid-therapy is sessions five to eight, and late in therapy is session nine onwards. Studies were not separated out in terms of presenting problems. This is due to the high levels of comorbidity between anxiety and depression (Kessler et al., 2006) and the relatively low numbers of papers looking at anxiety only (n = 4). Due to the review considering several factors, some studies are considered under several headings. However, findings from the studies addressing the reverse causality hypothesis are considered separately, so as not to examine the same data twice.

The quality appraisal, methodological approach and results were then collated onto a database (see Tables one and two). The synthesis adopted a narrative approach as several studies included in the review did not publish sufficient data for meta-analysis to proceed.

Results

Figure 1 presents a summary of the review process. The search generated 1069 results, with screening resulting in 57 full text papers being obtained for further scrutiny. Following further screening of the full text against the inclusion criteria resulted in 21 studies were eligible for inclusion. Key details of the studies are summarised in Table 1.

Strengths and Weaknesses of Included Studies

A summary of quality assessment scores is in Appendix C. Of the 21 papers assessed for quality, one (Tang & DeRubeis, 1999) scored considerably lower than others with a total score of six. The paper had poor external and internal validity and did not report its methodology clearly. This paper was excluded from the review, leaving 20 papers for data extraction. Amongst the remaining studies, quality was variable, with scores of 13-23. Only one of the studies reported a power analysis, with eight of the studies having a sample size of under 50. Small sample sizes increase the possibility of type II errors (Ellis, 2010). Several of the studies scored poorly for selection bias. Only five of the studies measured therapist competence or adherence to therapy techniques. Eight studies had both WA and outcome rated by clients which can increase the risk of Type I errors (Elvins & Green, 2008).

Study Characteristics

Table 1 shows the designs and critical appraisal of the studies in the review. Some of the studies had comparator groups that did not fit the inclusion and exclusion criteria Where this was the case, only the data and findings for the relevant group will be reported. The cumulative number of participants in the studies was 1539, with an average of 77 and a range of 19 to 367. Seven studies were randomised-controlled trials (RCTs), six took data from arms of RCTs and seven were cohort studies. Thirteen of the studies included participants experiencing depression, four included participants experiencing anxiety disorders and three included participants experiencing depression and/or anxiety. The interventions for included studies were as follows: eight CBT, nine cognitive therapy (CT) and two Cognitive Behavioural System of Psychotherapy (CBASP). One study compared CBT to Mindfulness Based Cognitive Therapy (MBCT), both of which were included in the review.

Table 1. *Methods and methodological issues*

Author (year)	Sample	Design	Presenting problem	Intervention	Strengths	Weaknesses
Arnow et al. (2013)	Full sample 224 (111 CBSAP + medication, 113 for BSP + medication)	RCT	MDD	CBASP or BSP plus antidepressant medication	Large sample size, therapist adherence measures used, analysis controlled for prior symptom improvement.	Outcome and WA both rated by client, didn't examine development of alliance-outcome relationship across therapy.
DeRubeis & Feeley (1990)	25	Cohort	MDD	CT	Examines temporal relationship of WA and outcome.	Small sample size, WA rated by undergraduates, low IRR for WA ratings.
Feeley, DeRubeis, & Gelfand (1999)	30	RCT	MDD	CT	Naturalistic setting, controlled for prior symptom change in analyses.	Small sample size, WA rated by undergraduates, low IRR for WA ratings.
Hardy et al. (2001)	24	Cohort	MDD	CT	Naturalistic setting, sessional measures of WA.	Small sample size, outcome and WA measured both rated by client, no measurement of therapist adherence to treatment.
Haug et al. (2016)	82	2 arms from RCT (adults face to face or online CBT)	SA & PD	CBT	Large sample size, naturalistic setting, good competence measures for therapists, controlled for prior symptom change in analyses.	Same raters for therapy competence and adherence (confounds), measure for competence and adherence not validated, large amounts of missing data.

Table 1. *Methods and methodological issues*

Author (year)	Sample	Design	Presenting problem	Intervention	Strengths	Weaknesses
Huppert et al. (2013)	133	2 arms from RCT (CBT + or - medication)	Panic disorder with agoraphobia	CBT	Large sample size, analyses controlled for prior symptom improvement, separated out therapist and client contributions to the WA.	WA and outcome measured (in part) both by clients, low number of WA ratings.
Klein et al. (2003)	367	1 arm of RCT	MDD	CBASP with or without medication	Large sample size, examined temporal effects of WA on outcome, used multiple statistical measures, used therapist adherence measures, controlled for prior symptom change in analyses.	
Lorenzo-Luaces, DeRubeis & Webb (2014)	60	1 arm of RCT	MDD	CT	Relatively large sample size, controlled for prior symptom change in analyses.	No therapist adherence or competence measures.
Marmar, Gaston, Gallagher, & Thompson (1989)	Full sample 60 (CT 22, BT 16, BDT 22)	RCT	MDD	CT, BT or brief dynamic therapy.	Good description of therapist adherence measures, therapists trained to use WA measure, use of both client and therapist measured WA.	WA only measured at one point during therapy, very small sample of CT clients for analysis.
Muran et al. (1995)	53	Cohort	Depression or anxiety	CT	Naturalistic setting.	Outcome and WA both measured by client, missing data, used mean WA score rather than session by session.

Table 1. *Methods and methodological issues*

Author (year)	Sample	Design	Presenting problem	Intervention	Strengths	Weaknesses
Preschel, Maercker & Wagner (2011)	Full sample 62 (32 group, 30 individual)	RCT	Depression	CBT	Controlled for initial symptom improvement in the analysis.	Low threshold on BDI for inclusion.
Saatsi, Hardy & Cahill (2007)	110	Cohort	Depression	CBT	Use of clinical interview to determine depression, naturalistic setting.	Two different WA measures used across sample, no measure of treatment competence or adherence, outcome and WA both rated by client.
Safran & Wallner (1991)	22	Cohort	Depression or anxiety	CT augmented - emphasis on WA	Use of two WA measures.	WA only measured at one point during therapy, no control for initial symptom improvement, small sample size, no therapist adherence or competence measures, outcome and WA both rated by client.
Snippe et al. (2015)	75	RCT	Depression	CBT or MBCT	Analyses controlled for prior symptom improvement, relatively large sample size.	Low threshold for inclusion, WA and outcome both rated by clients.
Strunk, Cooper, Ryan, DeRubeis & Hollon (2012)	176	1 arm of RCT	MDD	CT plus anti-depressant medication	Good training for WA raters and each therapy session rated by four raters, controlled for prior symptom change in analyses.	Only looks at WA in the early stages of therapy, no therapist adherence or competence measures.

Key features and findings of the studies are presented in table two. Nine studies used the Working Alliance Inventory (WAI; Horvath & Greenberg, 1986, 1989); four used the

Table 1. *Methods and methodological issues*

Author (year)	Sample	Design	Presenting problem	Intervention	Strengths	Weaknesses
Trepka, Rees, Shapiro, Hardy & Barkham (2004)	30	Cohort	Depression	CT	Good follow up (four months), naturalistic setting.	Use of overall mean alliance score, WA and outcome both rated by clients, small sample size.
Weck et al. (2016)	59	RCT	PD with agoraphobia	CBT (in vivo exposure with or without interoceptive exposure)	Used measures of therapist competence and adherence, looks at follow-up.	WA only measured at one point during therapy, results at end of therapy not reported.
Weck, Grikscheit, Jakob, Hofling & Stangier (2015)	61	1 arm from 3 RCT's aggregated	Depression or anxiety	CBT	Looks at temporal relationship of WA and outcome.	Heterogeneous sample as taken from three clinical trials, WA and outcome both rated by clients.
Weiss, Kivity & Huppert (2014)	19	Cohort	PD	CBT	Use of sessional WA measures.	Small sample size, inexperienced therapists and no competence or adherence measures used.
Zuroff & Blatt (2006)	99 full sample (CBT 50, IPT 49)	RCT	MDD	CBT or IPT	Good inter-rater reliability for WA raters, analysis controlled for prior symptom improvement.	Only used one sub-scale of alliance measure.

Note. CBASP = Cognitive Behavioural System of Psychotherapy; BSP = Brief Supportive Psychotherapy; CBT = Cognitive Behaviour Therapy; PA = Psychoanalytic Psychotherapy; PD = Psychodynamic Psychotherapy; CT = Cognitive Therapy; BT = Brief Therapy; BDT = Brief Dynamic Therapy; IPT = Interpersonal Therapy; RCT = Randomised Controlled Trial; MDD = Major Depressive Disorder; SA = Social Anxiety; PD = Panic Disorder; MBCT = Mindfulness Based Cognitive Therapy

Helping Alliance Questionnaire (HAQ; Luborsky, 1984), three used the California Alliance Scales (CALPAS; Marmar & Gaston, 1989); two used the CALPAS and the Agnew Relationship Measure (ARM; Agnew-Davies, Hardy, Barkham, Stiles & Shapiro, 1998); one used the CALPAS and the WAI and one study used the Barratt-Lennard Relationship Inventory (BLRI; Barrett- Lennard, 1986) and the patient contribution of the Vanderbilt Therapy Alliance Scales (VTAS; Krupnick et al., 1996). Ten of the studies used client rated alliance, six used observer rated alliance, three used client and therapist rated alliance and one used client and observer rated alliance. None of the studies used therapist ratings of alliance only.

Mean Ratings of Alliance and Outcome

Four of the included studies looked at the alliance-outcome correlation by aggregating a mean score gathered from sessional outcome data. All four of the studies found a significant correlation between WA and outcome (where higher alliance was associated with reduced symptomology), with effect sizes ranging from small to large (r = .29 to r = .50; Cohen, 1988). Both Hardy et al. (2001) and Saatsi et al (2007) found that the impact of a client's interpersonal style was mediated by their WA. While Trepka et al. (2004) found that the alliance-outcome correlation was present for both therapy completers and non-completers, Saatsi, Hardy and Cahill (1997) found that, when the sample was divided, the alliance-outcome correlation was only significant for therapy completers. While all four of the studies benefited from being conducted in a naturalistic clinical setting, only one (Saatsi, Hardy & Cahill, 1997) controlled for symptom severity at intake, meaning that symptom improvement prior to alliance measurement may have confounded outcome.

Early Alliance and Outcome

Ten of the studies examined the alliance-outcome correlation where alliance was measured early in treatment. Effect sizes for the alliance-outcome correlation ranged from small to medium (r = <.1 to .45) with nine studies reporting some significant findings for the alliance-outcome relationship. Four studies had mixed findings. Snippe et al. (2015) compared CBT to MBCT and found no significant relationship between alliance and outcome for MBCT, with small effect sizes. Alliance significantly predicted outcome with a medium effect size for CBT at session four. Strunk et al. (2012), found that when they controlled for prior symptom change and medication use in their regression model, the alliance-outcome relationship was no longer statistically significant with a small effect size. This suggests that the statistical approach and controlling for prior change may impact on findings. This is illustrated by the findings of Preschel, Maerker & Wagner (2011) who found a significant alliance-outcome correlation for client but not therapist rated alliance measured early and late in therapy, reporting a small effect size, but not for the alliance and the residual gain score. Weiss, Kivity & Huppert (2014) found that early alliance did not predict outcome, however the authors concluded that this was due to the sample being underpowered as the medium effect sizes obtained fall in line with other research.

Six studies had broadly positive findings for the alliance-outcome relationship. Arnow et al. (2013) found that early alliance was significantly associated with subsequent symptom reduction, reporting a small effect size. Klein et al. (2003) employed a large sample size and found that WA still predicted outcome while controlling for several client variables in two separate statistical models, with small effect sizes. Huppert et al. (2012) used multi-level modelling to investigate therapist and client contributions to the alliance. They found that client contributions to the alliance at session three significantly predicted outcome with medium effect sizes while therapist contributions only significantly predicted outcome for a limited range of measures.

Lorenzo-Luaces, DeRubeis & Webb (2014) found that alliance was significantly associated with outcome in their full sample, reporting small effect sizes, this was moderated by the number of prior episodes of depression. When the sample was divided, the alliance-outcome relationship only remained for those with fewer episodes of depression. In the only study that compared WA measures, Safran & Wallner (1991) found a significant alliance-outcome relationship for the CALPAS with medium effect sizes but not the WAI, reporting small to medium effect sizes. Neither measure predicted outcome for a general symptoms outcome measure. Weck et al. (2015) found that alliance scores early in treatment (including some from session five) were significantly higher for clients considered 'treatment successes' than 'treatment failures,' however outcomes were an aggregation of different measures due to being taken from different clinical trials and the authors did not control for prior symptom improvement in their analyses.

Table 2. Summary of findings

Author (year)	Point at which WA measured	Who completed WA measure	WA measure	Outcome measures	Findings	ES
Arnow et al. (2013)	Early (wk. 2 or 4), middle (wk. 6) or late (week 12)	Client	WAI-S	HAM-D	Early alliance predicted reduction in symptoms across therapy, accounting for 4.4-12.3% of the total variance in depressive symptoms between weeks 4 and 12.	r = .45
DeRubeis & Feeley (1990)	Second session (early) plus one randomly selected session from quadrant 2 (weeks 4-7), quadrant 3 (weeks 7-9), and quadrant 4 (weeks 10-12)	Observer	HAQ	BDI	No significant relationship between WA and outcome (either prior or subsequent change) for early session or quadrants 2 or 3. WA was related to prior change outcome at quadrant 4.	Early session: prior change r = .28/subsequent change r = .10; quadrant 2: prior change r = .02. subsequent change r = .14; quadrant 3 prior change r = .53; quadrant 4 prior change r = .42
Feeley, DeRubeis, & Gelfand (1999)	Session 2 plus one randomly selected session from quadrant 3 (weeks 7-9) and quadrant 4 (weeks 10-12).	Observer	HAQ	BDI	No significant relationship between WA and prior or subsequent symptom stage at session 2 or quadrants 3 or 4.	Session 2 subsequent change r = .27; quadrant 3 prior change r = .39; quadrant 4 prior change r = .15
Hardy et al. (2001)	After each session	Client	CALPAS	BDI	Alliance significantly predicted outcome over pretherapy status on BDI and in-treatment improvement. Alliance mediated the impact of under involved client interpersonal style.	r = .61
Haug et al. (2016)	After sessions 3 and 8	Client	WAI-S	BSQ for PD, SPS for SAD: scores were statistically transformed into a combined score.	WA not significantly associated with prior symptom improvement. There was a relationship between WA and outcome for session 8 but not session 3 of therapy.	session 3: r = .20; session 8: r = .37

Table 2. Summary of findings

Author (year)	Point at which WA measured	Who completed WA measure	WA measure	Outcome measures	Findings	ES
Huppert et al., 2014	After sessions 3 and 9	Client and Therapist	WAI-S	PDSS - both independent evaluator and self-report forms), ASI	Client contributions to WA were predictive of post-treatment outcome at session 3 and 9. Therapist contributions to WA were only predictive of outcome a limited number of measures and subscales.	r = .38
Klein et al. (2003)	after session 3/4 (wk. 2), after sessions 8-12 (wk. 6) and after session 16-20 (wk. 12)	Client	WAI-S	HRSD	Early alliance significantly predicted subsequent symptom change after controlling for prior change in depressive symptoms, gender, chronicity, comorbid diagnosis, substance use, PD, social functioning, history of abuse and neglect.	wk. 2: *r* = .24; wk. 6: *r* = .29; wk. 12: *r* = .40
Lorenzo-Luaces, DeRubeis & Webb (2014)	early (wk. 2, 3 or 4)	Observer	WAI-OS	BDI	WA was related to subsequent symptom change. Number of depressive moderated outcome.	*r* = .23
Marmar, Gaston, Gallagher, & Thompson (1989)	5th	Therapist & client	CALPAS	BDI	WA overall not related to outcome for CT, though 'patient commitment' scale was significantly related to outcome.	Overall therapist rated: *r* = .17; patient commitment: *r* = .73; patient working capacity: *r* = .34
Muran et al. (1995)	every session (post)	Client	CALPAS	SCL-90, IIP, ATQ, GAS, TC all at pre and post treatment	Significant relationship between WA and outcome found on GAS, TTC, PSC and TSC overall.	*r* = .29

Table 2. Summary of findings

Author (year)	Point at which WA measured	Who completed WA measure	WA measure	Outcome measures	Findings	ES
Preschel, Maercker & Wagner (2011)	4 (clients only) and post-treatment (client & therapist)	Client and Therapist	WAI	BDI	Significant relationship between client (but not therapist) rated WA and outcome at early and post treatment. WA scores did not significantly predict the BDI residual gain score at mid or post treatment.	session 4 client rated r = .40; post treatment client rated r = .42; post treatment therapist rated r = .24
Saatsi, Hardy & Cahill (2007)	Sessional	Client	CALPAS & Agnew Relationship Measure	BDI	Alliance was correlated with outcome. Alliance outcome correlation not sig for secure group or non-completers. The association between interpersonal style and outcome was mediated by the alliance.	r = .47; task r = .35, bond = .08, goal r = .33
Safran & Wallner (1991)	After 3rd session	Client	CALPAS, WAI	SCL-90, MCMI, BDI,	No significant correlations between WA and outcome on the anxiety subscale of the MCMI or the SCL-90. CALPAS but not WAI predicted outcome on the BDI and depression scale of the MCMI (2)	CALPAS-BDI: r = .45; CALPAS-MCMI: r = .45; WAI-BDI: r = .34; WAI-MCMI: r = .25
Snippe et al. (2015)	After sessions 2 and 4	Client	WAI	BDI	CBT: symptom improvement predicted by higher WAI scores as well as higher ratings of the Task and Bond subscales after session 4 while controlling for pre-treatment BDI scores. Alliance ratings after session 2 not associated with symptom improvement except for the Task subscale. No significant results found for MBCT for alliance.	CBT: r = .2 MCBT: r = < .1

Table 2. Summary of findings

Author (year)	Point at which WA measured	Who completed WA measure	WA measure	Outcome measures	Findings	ES
Strunk, Cooper, Ryan, DeRubeis & Hollon (2012)	Within first three sessions	Observer	WAI-S	BDI	Alliance predicted session to session symptom change. Only the task subscale of the WAI predicted S2S symptom change when subscales examined. When model accounted for medication and prior symptom change, alliance no longer significantly predicted outcome (trend level).	r = .16; model with medication and prior symptom change r = .14
Trepka, Rees, Shapiro, Hardy & Barkham (2004)	Sessional	Client	CALPAS & Agnew Relationship Measure	BDI: Intake, then sessional. Follow up - four months after treatment	Alliance ratings were associated with outcome for full sample and completers and non-completers at end of therapy and follow-up. Competence-outcome correlation was not independent of alliance; alliance -outcome correlation was independent of competence.	*r* = .5
Weck et al. (2016)	6	Observer	HAQ	BSI, PAS, AFPIB	Alliance made moderate to large effects on change in agoraphobia avoidance behaviour at follow-up, and was positively associated with outcome at follow up	*r* = .31
Weck, Grikscheit, Jakob, Hofling & Stangier (2015)	Three therapy sessions from session 1-5 (mostly 1-3)	Observer	HAQ	HRSD, BDI, LSAS, SPIN, Y-BOCS, IAS	Significantly higher scores on WA in treatment successes. WA was a significant moderator of adherence-outcome relationship. WA was a significant mediator for competence-outcome relationship. Competence and alliance were significant mediators for adherence-outcome relationship.	*r* = .33

Table 2. Summary of findings

Author (year)	Point at which WA measured	Who completed WA measure	WA measure	Outcome measures	Findings	ES
Weiss, Kivity & Huppert (2014)	Before and after each session	Client	WAI-S	PDSS-SR, ASI	Early alliance did not significantly predict t decreases in ASI scores during therapy. A sawtooth pattern of alliance was observed.	R^2= 13.99%
Zuroff & Blatt (2006)	3	Client, observer	VTAS (patient contribution only)	BDI, HRSD, SCL-90, GAS, SAC turned into a composite	WA predicted outcome in CBT	r = .21

Note. WAI = Working Alliance Inventory; WAI-S = Working Alliance Inventory (short version); HAQ = Penn Helping Alliance Scale; CALPAS = California Psychotherapy Alliance Scales; WAI-O = Working Alliance Inventory (observer version); B-L RI = Barrett-Lennard Relationship Inventory; HAM-D = Hamilton Depression Scale; BDI = Beck Depression Inventory; SCL-90 = Symptoms Checklist-90; Hamilton Depression Rating Scale (HDRS); GAS = Global Assessment of Symptoms; BSQ = Body Sensation Questionnaire; SPS = Social Phobia Scale; PDSS = Panic Disorder Symptoms Scale; ASI = Anxiety Sensitivity Index; IIP = Inventory of Interpersonal Problems; ATQ = Automatic Thoughts Questionnaire; TC = Target Complaints; MCMI = Millon Multiaxial Clinical Inventory; SAC = Social Adjustment Scale; EAC = Enhanced Capacities; BSI = Brief Symptom Inventory; PAS = Panic and Agoraphobia Scale; AFIB = Assessment Form of Interpersonal Behaviour; LSAS = Liebowitz Social Anxiety Scale; SPIN = Social Phobia and Anxiety Inventory; Y-BOCS = Yale-Brown Obsessive Compulsive Scale for Hypochrondriasis; IAS = Illness Attitude Scales

Mid Therapy Alliance and Outcome

Three studies examined the alliance-outcome relationship at mid-therapy. Effect sizes ranged from small to large (r = .17 – r = .74), two of which were statistically significant. Klein et al. (2003) employed a large sample size and found a small but significant effect size for the alliance-outcome relationship. Marmar et al. (1989) found no significant relationship between therapist rated WA subsequent treatment outcome, even though they had trained therapists in rating alliance. However, client rated commitment and working capacity were significantly related to outcome with large and medium effect sizes respectively. Weck et al. (2015) found a significant relationship between WA at session 6 and outcome at follow up with an effect size of 0.31. Although Weck et al. (2015) were amongst the few studies to complete follow up analyses, they did not report findings for alliance and outcome at the end of therapy.

Late Alliance and Outcome

Three studies examined the alliance-outcome relationship in late therapy (Huppert et al., 2014; Klein et al., 2003; Preschel, Maerker & Wagner, 2011). Effect sizes were all significant and in the medium range (r = .38 - .42).

Subscales of Alliance Measures and Outcome

Six studies examined how alliance measure subscales were associated with outcome. Three used the WAI, two used the CALPAS and one used the patient contribution of the VTAS. For the WAI, two studies (Huppert et al., 2014; Snippe et al., 2015) found significant relationships between the task and bond subscales and outcome but not the goal subscale, while the third study found that the goal and task subscales but not the bond subscale were significant (Saatsi, Hardy & Cahill, 2007). For the CALPAS, Marmar et al. (1989) found that patient commitment and patient working capacity were significantly related to outcome with large and medium effect sizes respectively. Safran and Wallner (1991) similarly found that Patient Commitment,

Patient Working Capacity and Goal disagreement were significantly related to outcome, with the remaining therapist subscales yielding small to medium effect sizes. Zuroff & Blatt (2006) found that the patient contribution of the VTAS was significantly correlated with outcome with a small effect size. However, correlations with other aspects of the VTAS were not reported.

Reverse Causality Hypothesis

Three studies examined whether symptom improvement predicted changes in the WA or vice versa. Findings were mixed. De Rubeis and Feeley (1990) found no alliance-outcome correlation at early, mid or late therapy but did find that WA at the end of therapy was significantly associated with prior symptom change. When replicating this study (Feeley, DeRubeis and Gelfland, 1999) the authors found no significant relationship between alliance and prior or subsequent symptom change. However, the effect size for subsequent change at session two was small and for prior change at quadrant three was medium. The authors did not examine alliance and subsequent change (only prior) after session seven of therapy, thus limiting the potential findings. Haug et al. (2016) employed a larger sample size and found that WA scores were not significantly associated with prior symptom improvement. They were significantly associated with subsequent symptom improvement when alliance was measured at session eight but not three of CBT with medium and small effect sizes respectively.

Discussion

The purpose of this systematic literature review was to examine the relationship between WA and outcome in CBT for anxiety and depression, with a focus on the temporal relationship between alliance and outcome and how the underlying features of alliance relate to outcome. Some tentative findings regarding the alliance-outcome relationship for CBT for anxiety and depression can be made. Overall, there is broad

support for the alliance-outcome relationship measured at any stage of therapy. This is in line with previous meta-analyses such as Flückiger et al. (2012) who found an overall effect size of .29 and that the alliance-outcome relationship was not moderated by CBT. The effect sizes for the alliance-outcome relationship within this review ranged from small to large (from <0.1 to 0.61) suggesting a broad range of findings within the studies. This is in line with previous findings of heterogeneity in alliance-outcome data (Horvath & Bedi, 2002; Horvath & Symonds, 1991, 2011; Martin, Garske & Davis, 2000).

There is some evidence that the approaches used to investigate the alliance-outcome relationship influence findings. Using a mean score from multiple sessions across therapy appeared to consistently produce positive findings for the alliance-outcome relationship in this review. This is broadly in line with the finding of Horvath, Del Re, Flückiger & Symonds (2011) that effect sizes for the alliance-outcome correlation when mean ratings of alliance were used were higher than early or middle assessments of alliance. However, this method may obscure patterns in alliance development (Doran, 2014). There was further indication from other studies that the methodology used impacted on the outcome of studies. The two studies that investigated therapist rated alliance and outcome (in addition to client rated alliance) did not have significant findings. This is in line with previous research that has found therapist rated alliance to be least predictive of outcome (Horvath & Symonds, 1991).

There was some evidence to suggest that the alliance outcome relationship may not be as important for some variants of CBT than others, with very small effect sizes found for MBCT and larger effect sizes found for CBASP. This supports the suggestion that alliance may be more important for different therapies (e.g. Gaston, Thompson, Gallagher, Cournoyer & Gagnon, 1998) though it would be unwise to draw conclusions based on such a small number of studies. There is some evidence of factors that

additionally influence the alliance-outcome relationship, such as prior number of episodes of depression, interpersonal style, social functioning, gender and chronicity of difficulty. Thus, the alliance-outcome relationship is complex and the variables that contribute to or confound it are not fully understood (Doran, 2014).

There was insufficient evidence to suggest whether the alliance-outcome relationship changes over the course of CBT for anxiety and depression. This was in part because many studies examined alliance early in therapy, with fewer looking at alliance at the mid and late stages of therapy. However, for the three studies that examined alliance and subsequent outcome at several times points, all found that effect sizes for the alliance-outcome correlation increased over the course of therapy (Haug et al., 2016; Klein et al., 2003; Preschel, Maerker & Wagner, 2003). This falls in line with cross therapy meta-analyses (Horvath, Del Re, Flückiger and Symonds 2011) that found that the effect sizes for the alliance-outcome correlation significantly increased as therapy progressed.

In terms of aspects of the alliance that might be important for CBT, some tentative findings can be made. There was evidence from several studies that the task subscale of the WAI is important in CBT, though there were conflicting findings as to the importance of the goal and bond subscales. There was some limited evidence from the CALPAS that client contributions to the alliance might be more important than therapist contributions. However, these findings should be regarded with caution, as they are based on a limited number of studies and there is evidence that the subscales of alliance inventories are often highly correlated (Elvins & Green, 2008).

The findings in terms of the reverse causality hypothesis were unclear. While there was some evidence that alliance sometimes predicts prior symptom improvement, particularly later in therapy, this was not a consistent finding. This goes some way to supporting the assertion of Crits-Christoph, Gibbons, Hamilton, Ring-Kurtz & Gallop

(2011) who suggest that later in therapy, symptom improvement predicts alliance, and the alliance-outcome relationship is therefore better examined earlier in therapy. The number of studies that controlled for prior improvement and still found a significant alliance-outcome relationship would run counter to claims that alliance is purely caused by symptom improvement (DeRubeis & Feely, 1990); however, controlling for prior symptom change does decrease effect sizes (Huppert et al., 2014; Preschel, Maerker & Wagner, 2011).

Future Research

This review has highlighted a number of issues for future research to address. Firstly, it is important that future research uses appropriate methodology. Statistical controls to account for improvement prior to alliance measurement are essential, as well as the use of multiple alliance measurements. Client or observer ratings of alliance appear to be more predictive of outcome than therapist measures.

Secondly, there are unanswered questions about additional factors that may influence the alliance-outcome relationship. While there is some suggestion that client factors may impact on alliance (Hardy et al., 2001; Klein et al., 2003; Lorenzo-Luaces, DeRubeis & Webb, 2014; Saatsi, Hardy & Cahill, 2007), there is insufficient evidence to make any definitive conclusions about this and would benefit from further research.

Thirdly, though an alliance-outcome relationship appears to be present for CBT for anxiety and depression, the underlying mechanisms of change are unclear. There was insufficient evidence from the subscales of the measures used that any part of the alliance is more important. It is not possible to say, therefore whether CBT for anxiety and depression differs in terms of alliance compared to other therapies for the same presenting problems. Further research comparing the alliance for differing therapies could address this question.

Implications for Practice

This review suggests that, in line with previous research, therapist ratings of alliance are least predictive of CBT outcome (Horvath & Symonds, 1991). Given that alliance makes a significant contribution to outcome in CBT for anxiety and depression, it is important that therapists elicit feedback from clients about their experience of the alliance, possibly via client ratings of this relationship. In addition, given that client contributions to the alliance may be more important to outcome than therapist contributions, it may be beneficial to consider how aspects of the client's history and relationship style may impact the alliance in therapy. This is of relevance given evidence that the impact of interpersonal style on outcome is mediated by the alliance (Hardy et al., 2001; Saatsi et al., 2007). As alliance may have differential impact on outcome throughout the course of therapy, it is important to monitor the WA throughout therapy. In addition, the use of WA measures as a routine monitoring tool would allow for the collection of data to be used in further research.

Strengths and Limitations of the Review

The review is the first to address the alliance-outcome relationship specifically for CBT for anxiety and depression. The systematic approach to the literature search and subsequent appraisal of the literature reduced systemic bias. The review also has some limitations. The search was limited to English Language papers only, and no search of citations or grey literature was undertaken. Including only English language papers and not completing a citation search may have resulted in important papers not being included. Not searching the grey (unpublished) literature may have resulted in a bias towards the inclusion of significant findings in the review.

Some papers had to be excluded due to their looking at multiple groups, some of which were not relevant to the review. Only including those which published findings separately may have biased the review to papers who found significant findings for those groups. Because the review was investigating several aspects of the alliance,

findings from several studies were repeated in different sections of the review, potentially biasing the review towards those findings. However, overall, the review contributes to the alliance-outcome debate.

Conclusion

This systematic literature review provides support for the presence of the alliance-outcome relationship for CBT in anxiety and depression. Some tentative conclusions about the temporal relationship of alliance and outcome, important aspects of alliance and the importance of alliance for different variants of CBT for anxiety and depression can be made. However, the findings are limited due to methodological weaknesses of the review, including the limited number of studies and the weaknesses of the studies themselves, including sample size. Further research examining how client characteristics influence alliance and comparing different therapies are required to further understand the alliance-outcome relationship.

References

Agnew-Davies, R., Stiles, W.B., Hardy, G. E, Barkham, M., & Shapiro, D. A. (1998). Alliance structure assessed by the Agnew Relationship Measure (ARM). *British Journal of Clinical Psychology, 37,* 155-172. doi: 10.1111/j.2044-8260.1998.tb01291.x

*Arnow, B. A., Steidtmann, D., Blasey, C., Manber, R., Constantino, M. J., Klein, D. N., . . . Kocsis, J. H. (2013). The relationship between the therapeutic alliance and treatment outcome in two distinct psychotherapies for chronic depression. *Journal of Consulting and Clinical Psychology, 81*, 627-638. doi: 10.1037/a0031530

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed., rev.). Washington, DC: American Psychiatric Association.

Barber, J. P., Luborsky, L., Crits-Christoph, P., Thase, M. E., Weiss, R., Frank, A., . . . Gallop, R. (1999). Therapeutic alliance as a predictor of outcome in treatment of cocaine dependence. *Psychotherapy Res., 9*, 54-73. doi: 10.1093/ptr/9.1.54

Barrett-Lennard, G. T. (1986). The relationship inventory now: issues and advances in theory, method and use. In L. S. Greenberg & W.M. Pinsof (Eds.), *The psychotherapeutic process: A research handbook* (pp. 439−467). New York and London: Guilford Press.

Beck, A., Rush, A., Shaw, B., & Emery, B. (1979). The therapeutic relationship: application to cognitive therapy. In A. Beck, A. Rush, B. Shaw, & G. Emery, *Cognitive therapy of depression.* New York, NY: Guildford Press.

Beck, A. T. (1970). Cognitive therapy: Nature and relation to behavior therapy. *Behavior Therapy, 1*(2), 184-200. doi:10.1016/S0005-7894(70)80030-2

Bell, E. C., Marcus, D. K., & Goodlad, J. K. (2013). Are the parts as good as the whole? A meta-analysis of component treatment studies. *Journal of Consulting and Clinical Psychology, 81*, 722-736. doi:10.1037/a0033004

Borenstein, M. (2009). *Introduction to meta-analysis*. Chichester: John Wiley & Sons.

Castonguay, L. G., Constantino, M. J., & Holtforth, M. G. (2006). The working alliance: Where are we and where should we go? *Psychotherapy: Theory, Research, Practice, Training, 43*, 271-279. doi:10.1037/0033-3204.43.3.271

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6*, 284-290. doi:10.1037/1040-3590.6.4.284

Crits-Christoph, P., Gibbons, M. B. C., Hamilton, J., Ring-Kurtz, S., & Gallop, R. (2011). The dependability of alliance assessments: The alliance-outcome correlation is larger than you might think. *Journal of Consulting and Clinical Psychology, 79*, 267-278. doi:10.1037/a0023668

Das-Munshi, J., Goldberg, D., Bhugra, D. K., Dewey, M. E., Jenkins, R., Stewart, R., . . . Brugha, T. S. (2008). Public health significance of mixed anxiety and depression: Beyond current classification. *British Journal of Psychiatry, 192*, 171-177. doi:10.1192/bjp.bp.107.036707

*DeRubeis, R. J., & Feeley, M. (1990). Determinants of change in cognitive therapy for depression. *Cognitive Therapy and Research, 14*, 469-482. https://doi.org/10.1007/bf01172968

Doran, J. M. (2014). The working alliance: Where have we been, where are we going? *Psychotherapy Research,* 1-18. doi:10.1080/10503307.2014.954153

Downs, S. H., & Black, N. (1998). The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *Journal of Epidemiology and Community Health, 52*, 377. doi:10.1136/jech.52.6.377

Ellis, P. D. (2010). *The essential guide to effect sizes [electronic resource]: Statistical power, meta-analysis, and the interpretation of research results*. Cambridge: Cambridge University Press, 2010.

Elvins, R., & Green, J. (2008). The conceptualization and measurement of therapeutic alliance: An empirical review. *Clinical Psychology Review, 28*, 1167-1187. doi:10.1016/j.cpr.2008.04.002

*Feeley, M., DeRubeis, R. J., & Gelfand, L. A. (1999). The temporal relation of adherence and alliance to symptom change in cognitive therapy for depression. *Journal of consulting and clinical psychology, 67*, 578-582. doi: 10.1037//0022-006x.67.4.578

Flückiger, C., Del Re, A. C., Wampold, B. E., Symonds, D., & Horvath, A. O. (2012). How central is the alliance in psychotherapy? A multilevel longitudinal meta-analysis. *Journal of Counseling Psychology, 59*, 10-17. doi:10.1037/a0025749

Gaudiano, B. A. (2008). Cognitive-behavioural therapies: achievements and challenges. *Evidence Based Mental Health, 11*, 5. doi:10.1136/ebmh.11.1.5

Gaston, L., Thompson, L., Gallagher, D., Cournoyer, L.-G., & Gagnon, R. (1998). Alliance, technique, and their interactions in predicting outcome of behavioral, cognitive, and brief dynamic therapy. *Psychotherapy Research, 8*, 190-209. doi:10.1080/10503309812331332307

Gelso, C. J., & Hayes, J. A. (1998). *The psychotherapy relationship: Theory, research, and practice.* Hoboken, NJ: John Wiley.

*Hardy, G. E., Cahill, J., Shapiro, D. A., Barkham, M., Rees, A., & Macaskill, N. (2001). Client interpersonal and cognitive styles as predictors of response to time-limited cognitive therapy for depression. *Journal of consulting and clinical psychology, 69*, 841-845. doi: 10.1037//0022-006x.69.5.841

*Haug, T., Nordgreen, T., Ost, L. G., Tangen, T., Kvale, G., Hovland, O. J., . . . Havik, O. E. (2016). Working alliance and competence as predictors of outcome in cognitive behavioral therapy for social anxiety and panic disorder in adults. *Behaviour Research and Therapy, 77*, 40-51. doi:10.1016/j.brat.2015.12.004

Hayes, S. C. (2004). Acceptance and commitment therapy, relational frame theory, and the third wave of behavioral and cognitive therapies. *Behavior Therapy, 35*, 639-665. doi:10.1016/S0005-7894(04)80013-3

Horvath, A. O., & Greenberg, L. S. (1989). Development and validation of the working alliance inventory. *Journal of Counseling Psychology, 36*, 223-233. doi:10.1037/0022-0167.36.2.223

Horvath, A. O., & Symonds, B. D. (1991). Relation between working alliance and outcome in psychotherapy: A meta-analysis. *Journal of Counseling Psychology, 38*, 139-149. doi: 10.1037//0022-0167.38.2.139

*Huppert, J. D., Kivity, Y., Barlow, D. H., Gorman, J. M., Shear, M. K., & Woods, S. W. (2014). Therapist effects and the outcome-alliance correlation in cognitive behavioral therapy for panic disorder with agoraphobia. *Behaviour Research and Therapy, 52*, 26-34. doi: 10.1016/j.brat.2013.11.001

Kable, A. K., Pich, J., & Maslin-Prothero, S. E. (2012). A structured approach to documenting a search strategy for publication: A 12 step guideline for authors. *Nurse Education Today, 32*, 878-886. doi:10.1016/j.nedt.2012.02.022

Kivlighan, D. M., & Shaughnessy, P. (1995). Analysis of the development of the working alliance using hierarchical linear modeling. *Journal of Counseling Psychology, 42*, 338-349. doi:10.1037/0022-0167.42.3.338

*Klein, D. N., Schwartz, J. E., Santiago, N. J., Vivian, D., Vocisano, C., Castonguay, L. G., . . . Keller, M. B. (2003). Therapeutic alliance in depression treatment:

Controlling for prior change and patient characteristics. *Journal of Consulting and Clinical Psychology, 71*, 997-1006. doi: 10.1037/0022-006x.71.6.997

Krupnick, J. L., Sotsky, S. M., Simmons, S., Moyer, J., Elkin, L., & Watkins, J. T. (1996). The role of the therapeutic alliance in psychotherapy and pharmacotherapy outcome: findings in the National Institute of Mental Health Treatment of Depression Collaborative Research Programme. *Journal of Consulting and Clinical Psychology, 64*, 532−539. doi:10.1037/0022-006X.64.3.532

Longmore, R. J., & Worrell, M. (2007). Do we need to challenge thoughts in cognitive behavior therapy? *Clinical Psychology Review, 27*, 173-187. doi: 10.1016/j.cpr.2006.08.001

*Lorenzo-Luaces, L., DeRubeis, R. J., & Webb, C. A. (2014). Client characteristics as moderators of the relation between the therapeutic alliance and outcome in cognitive therapy for depression. *Journal of Consulting and Clinical Psychology, 82*, 368-373. doi: 10.1037/a0035994

Marmar, C. R., Weiss, D. S., & Gaston, L. (1989). Towards the validation of the California Therapeutic Alliance Rating System. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 1*, 46−52. doi: 10.1037/1040-3590.1.1.46

*Marmar, C. R., Gaston, L., Gallagher, D., & Thompson, L. W. (1989). Alliance and outcome in late-life depression. *Journal of Nervous and Mental Disease, 177*, 464-472. doi: 10.1097/00005053-198908000-00003

Martin, D. J., Garske, J. P., & Davis, M. K. (2000). Relation of the therapeutic alliance with outcome and other variables: A meta-analytic review. *Journal of Consulting and Clinical Psychology, 68*, 438-450. doi:10.1037/0022-006X.68.3.438

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods, 1,* 30-46. doi:10.1037/1082-989X.1.1.30

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine, 151,* 264–269. doi:10.7326/0003-4819-151-4-200908180-

*Muran, J. C., Gorman, B. S., Safran, J. D., Twining, L., Samstag, L. W., & Winston, A. (1995). Linking in-session change to overall outcome in short-term cognitive therapy. *Journal of consulting and clinical psychology, 63,* 651-657. doi: 10.1037//0022-006x.63.4.651

National Institute for Clinical Excellence (2009). *Depression: the treatment and management of depression in adults (CG90)*. Retrieved 17.03.14 from: http://www.nice.org.uk/nicemedia/pdf/CG90NICEguideline.pdf

National Institute for Clinical Excellence (2011a). *Common mental health problems: identification and pathways to care (CG123)*. Retrieved 29.06.17 from: https://www.nice.org.uk/guidance/cg123/chapter/introduction

National Institute for Clinical Excellence (2011b). *Generalised anxiety disorder and panic disorder (with or without agoraphobia) in adults: Management in primary, secondary and community care (CG113).* Retrieved 17.03.14 from: http://www.nice.org.uk/nice media/live/ 13314/52601/52601.pdf

Norcross, J. (2002). *Psychotherapy relationships that work: Therapist contributions and responsiveness to patients*. New York, NY: Oxford University Press.

*Preschl, B., Maercker, A., & Wagner, B. (2011). The working alliance in a randomized controlled trial comparing online with face-to-face cognitive-behavioral therapy for depression. *Bmc Psychiatry, 11*. doi:10.1186/1471-244x-11-189

*Saatsi, S., Hardy, G. E., & Cahill, J. (2007). Predictors of outcome and completion status in cognitive therapy for depression. *Psychotherapy Research, 17*, 185-195. doi: 10.1080/10503300600779420

*Safran, J. D., & Wallner, L. K. (1991). The relative predictive validity of two therapeutic alliance measures in cognitive therapy. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 3*, 188-195. doi: 10.1037//1040-3590.3.2.188

Siev, J., Huppert, J., & Chambless, D.L. (2009). The dodo bird, treatment technique, and disseminating empirically supported treatments. *Association of advancement of behaviour therapy, 32*, 69-76.

*Snippe, E., Fleer, J., Tovote, K. A., Sanderman, R., Emmelkamp, P. M. G., & Schroevers, M. J. (2015). The therapeutic alliance predicts outcomes of cognitive behavior therapy but not of mindfulness-based cognitive therapy for depressive symptoms. *Psychotherapy and Psychosomatics, 84*, 314-315. doi:10.1159/000379755

*Strunk, D. R., Cooper, A. A., Ryan, E. T., DeRubeis, R. J., & Hollon, S. D. (2012). The process of change in cognitive therapy for depression when combined with antidepressant medication: Predictors of early intersession symptom gains. *Journal of Consulting and Clinical Psychology, 80*, 730-738. doi: 10.1037/a0029281

Tang, T. Z., & DeRubeis, R. J. (1999). Sudden gains and critical sessions in cognitive-behavioral therapy for depression. *Journal of Consulting and Clinical Psychology, 67*, 894-904. doi: 10.1037//0022-006x.67.6.894

*Trepka, C., Rees, A., Shapiro, D. A., Hardy, G. E., & Barkham, M. (2004). Therapist competence and outcome of cognitive therapy for depression. *Cognitive Therapy and Research, 28*, 143-157. doi: 10.1023/b:cotr.0000021536.39173.66

Centre for Reviews Dissemination (2009). *Systematic reviews: CRD's guidance for undertaking reviews in health care*. York: University of York.

Wampold, B. E. (2015). How important are the common factors in psychotherapy? An update. *World Psychiatry, 14,* 270-277. doi:10.1002/wps.20238

Wampold, B. E., Mondin, G. W., Moody, M., Stich, F., Benson, K., & Ahn, H. N. (1997). A meta-analysis of outcome studies comparing bona fide psychotherapies: Empirically, "all must have prizes". *Psychological Bulletin, 122,* 203-215. doi:10.1037/0033-2909.122.3.203

Webb, C. A., Derubeis, R. J., & Barber, J. P. (2010). Therapist adherence/competence and treatment outcome: A meta-analytic review. *Journal of Consulting and Clinical Psychology, 78,* 200-211. doi:10.1037/a0018912

*Weck, F., Grikscheit, F., Hofling, V., Kordt, A., Hamm, A. O., Gerlach, A. L., . . . Lang, T. (2016). The role of treatment delivery factors in exposure-based cognitive behavioral therapy for panic disorder with agoraphobia. *Journal of Anxiety Disorders, 42,* 10-18. doi:10.1016/j.janxdis.2016.05.007

*Weck, F., Grikscheit, F., Jakob, M., Hofling, V., & Stangier, U. (2015). Treatment failure in cognitive-behavioural therapy: Therapeutic alliance as a precondition for an adherent and competent implementation of techniques. *British Journal of Clinical Psychology, 54,* 91-108. doi:10.1111/bjc.12063

*Weiss, M., Kivity, Y., & Huppert, J. D. (2014). How does the therapeutic alliance develop throughout cognitive behavioral therapy for panic disorder? Sawtooth patterns, sudden gains, and stabilization. *Psychotherapy Research, 24,* 407-418. doi:10.1080/10503307.2013.868947

*Zuroff, D. C., Koestner, R., Moskowitz, D. S., McBride, C., Marshall, M., & Bagby, R. M. (2007). Autonomous motivation for therapy: A new common factor in

brief treatments for depression. *Psychotherapy Research, 17*, 137-148. doi:10.1080/10503300600919380

*Studies included in the review

Appendices

Appendix A: Search Strategy

Lit searches 01.03.17

(helping alliance OR therap* alliance OR working alliance OR therap* relationship) AND (cognitive behavio* therapy OR CBT) AND (outcome OR symptom reduction)

PsychInfo

1	helping alliance.mp.	386	Advanced
2	therap* alliance.mp.	6371	Advanced
3	working alliance.mp.	2821	Advanced
4	1 or 2 or 3	7967	Advanced
5	exp Cognitive Behavior Therapy/ or cognitive behavio* therapy.mp.	23982	Advanced
6	exp Cognitive Therapy/ or CBT.mp.	21789	Advanced
7	5 or 6	33718	Advanced
8	outcome.mp. or exp TREATMENT OUTCOMES/	167748	Advanced

9	symptom reduction.mp.	1517	Advanced
10	8 or 9	168761	Advanced
11	4 and 7 and 10	351	Advanced
12	therap* relationship.mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]	8116	Advanced
13	4 or 12	15245	Advanced
14	7 and 10 and 13	435	Advanced
15	limit 14 to english language	419	Advanced

419 articles retrieved

Medline

1	helping alliance.mp.	100	Advanced
2	therap* alliance.mp.	1965	Advanced
3	working alliance.mp.	654	Advanced
4	1 or 2 or 3	2497	Advanced

5	exp Cognitive Behavior Therapy/ or cognitive behavio* therapy.mp.	26255	Advanced
6	exp Cognitive Therapy/ or CBT.mp.	24873	Advanced
7	5 or 6	27960	Advanced
8	outcome.mp. or exp TREATMENT OUTCOMES/	1527100	Advanced
9	symptom reduction.mp.	1355	Advanced
10	8 or 9	1527776	Advanced
11	4 and 7 and 10	260	Advanced
12	therap* relationship.mp.	2210	Advanced
13	4 or 12	4536	Advanced
14	7 and 10 and 13	322	Advanced
15	limit 14 to english language	313	Advanced

313 articles retrieved

Web of Science (core collection)

You searched for: TOPIC: ("helping alliance" OR "therap* alliance" OR "working alliance" OR "therap* relationship") *AND* TOPIC: ("cognitive behavio* therapy" OR CBT) *AND* TOPIC: (outcome OR "symptom reduction")

Refined by: LANGUAGES: (ENGLISH)

Timespan: All years. Indexes: SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-EXPANDED, IC.

337 articles retrieved

Appendix B: Quality Assessment Tool (Downs & Black, 1998)

Appendix

Checklist for measuring study quality

Reporting

1. *Is the hypothesis/aim/objective of the study clearly described?*

yes	1
no	0

2. *Are the main outcomes to be measured clearly described in the Introduction or Methods section?*
   If the main outcomes are first mentioned in the Results section, the question should be answered no.

yes	1
no	0

3. *Are the characteristics of the patients included in the study clearly described ?*
   In cohort studies and trials, inclusion and/or exclusion criteria should be given. In case-control studies, a case-definition and the source for controls should be given.

yes	1
no	0

4. *Are the interventions of interest clearly described?*
   Treatments and placebo (where relevant) that are to be compared should be clearly described.

yes	1
no	0

5. *Are the distributions of principal confounders in each group of subjects to be compared clearly described?*
   A list of principal confounders is provided.

yes	2
partially	1
no	0

6. *Are the main findings of the study clearly described?*
   Simple outcome data (including denominators and numerators) should be reported for all major findings so that the reader can check the major analyses and conclusions. (This question does not cover statistical tests which are considered below).

yes	1
no	0

7. *Does the study provide estimates of the random variability in the data for the main outcomes?*
   In non normally distributed data the inter-quartile range of results should be reported. In normally distributed data the standard error, standard deviation or confidence intervals should be reported. If the distribution of the data is not described, it must be assumed that the estimates used were appropriate and the question should be answered yes.

yes	1
no	0

8. *Have all important adverse events that may be a consequence of the intervention been reported?*
   This should be answered yes if the study demonstrates that there was a comprehensive attempt to measure adverse events. (A list of possible adverse events is provided).

yes	1
no	0

9. *Have the characteristics of patients lost to follow-up been described?*
   This should be answered yes where there were no losses to follow-up or where losses to follow-up were so small that findings would be unaffected by their inclusion. This should be answered no where a study does not report the number of patients lost to follow-up.

yes	1
no	0

10. *Have actual probability values been reported(e.g. 0.035 rather than <0.05) for the main outcomes except where the probability value is less than 0.001?*

yes	1
no	0

External validity

All the following criteria attempt to address the representativeness of the findings of the study and whether they may be generalised to the population from which the study subjects were derived.

11. *Were the subjects asked to participate in the study representative of the entire population from which they were recruited?*
    The study must identify the source population for patients and describe how the patients were selected. Patients would be representative if they comprised the entire source population, an unselected sample of consecutive patients, or a random sample. Random sampling is only feasible where a list of all members of the relevant

population exists. Where a study does not report the proportion of the source population from which the patients are derived, the question should be answered as unable to determine.

yes	1
no	0
unable to determine	0

12. *Were those subjects who were prepared to participate representative of the entire population from which they were recruited?*
The proportion of those asked who agreed should be stated. Validation that the sample was representative would include demonstrating that the distribution of the main confounding factors was the same in the study sample and the source population.

yes	1
no	0
unable to determine	0

13. *Were the staff, places, and facilities where the patients were treated, representative of the treatment the majority of patients receive?*
For the question to be answered yes the study should demonstrate that the intervention was representative of that in use in the source population. The question should be answered no if, for example, the intervention was undertaken in a specialist centre unrepresentative of the hospitals most of the source population would attend.

yes	1
no	0
unable to determine	0

Internal validity - bias

14. *Was an attempt made to blind study subjects to the intervention they have received ?*
For studies where the patients would have no way of knowing which intervention they received, this should be answered yes.

yes	1
no	0
unable to determine	0

15. *Was an attempt made to blind those measuring the main outcomes of the intervention?*

yes	1
no	0
unable to determine	0

16. *If any of the results of the study were based on "data dredging", was this made clear?*
Any analyses that had not been planned at the outset of the study should be clearly indicated. If no retrospective unplanned subgroup analyses were reported, then answer yes.

yes	1
no	0
unable to determine	0

17. *In trials and cohort studies, do the analyses adjust for different lengths of follow-up of patients, or in case-control studies, is the time period between the intervention and outcome the same for cases and controls ?*
Where follow-up was the same for all study patients the answer should yes. If different lengths of follow-up were adjusted for by, for example, survival analysis the answer should be yes. Studies where differences in follow-up are ignored should be answered no.

yes	1
no	0
unable to determine	0

18. *Were the statistical tests used to assess the main outcomes appropriate?*
The statistical techniques used must be appropriate to the data. For example non-parametric methods should be used for small sample sizes. Where little statistical analysis has been undertaken but where there is no evidence of bias, the question should be answered yes. If the distribution of the data (normal or not) is not described it must be assumed that the estimates used were appropriate and the question should be answered yes.

yes	1
no	0
unable to determine	0

19. *Was compliance with the intervention/s reliable?*
Where there was non compliance with the allocated treatment or where there was contamination of one group, the question should be answered no. For studies where the effect of any misclassification was likely to bias any association to the null, the question should be answered yes.

yes	1
no	0
unable to determine	0

20. *Were the main outcome measures used accurate (valid and reliable)?*

For studies where the outcome measures are clearly described, the question should be answered yes. For studies which refer to other work or that demonstrates the outcome measures are accurate, the question should be answered as yes.

yes	1
no	0
unable to determine	0

Internal validity - confounding (selection bias)

21. *Were the patients in different intervention groups (trials and cohort studies) or were the cases and controls (case-control studies) recruited from the same population?*
For example, patients for all comparison groups should be selected from the same hospital. The question should be answered unable to determine for cohort and case-control studies where there is no information concerning the source of patients included in the study.

yes	1
no	0
unable to determine	0

22. *Were study subjects in different intervention groups (trials and cohort studies) or were the cases and controls (case-control studies) recruited over the same period of time?*
For a study which does not specify the time period over which patients were recruited, the question should be answered as unable to determine.

yes	1
no	0
unable to determine	0

23. *Were study subjects randomised to intervention groups?*
Studies which state that subjects wererandomised should be answered yes except where method of randomisation would not ensure random allocation. For example alternate allocation would score no because it is predictable.

yes	1
no	0
unable to determine	0

24. *Was the randomised intervention assignment concealed from both patients and health care staff until recruitment was complete and irrevocable?*
All non-randomised studies should be answered no. If assignment was concealed from patients but not from staff, it should be answered no.

yes	1
no	0
unable to determine	0

25. *Was there adequate adjustment for confounding in the analyses from which the main findings were drawn?*
This question should be answered no for trials if: the main conclusions of the study were based on analyses of treatment rather than intention to treat; the distribution of known confounders in the different treatment groups was not described; or the distribution of known confounders differed between the treatment groups but was not taken into account in the analyses. In non-randomised studies if the effect of the main confounders was not investigated or confounding was demonstrated but no adjustment was made in the final analyses the question should be answered as no.

yes	1
no	0
unable to determine	0

26. *Were losses of patients to follow-up taken into account?*
If the numbers of patients lost to follow-up are not reported, the question should be answered as unable to determine. If the proportion lost to follow-up was too small to affect the main findings, the question should be answered yes.

yes	1
no	0
unable to determine	0

Power

27. *Did the study have sufficient power to detect a clinically important effect where the probability value for a difference being due to chance is less than 5%?*
Sample sizes have been calculated to detect a difference of x% and y%.

	Size of *smallest* intervention group	
A	$<n_1$	0
B	n_1–n_2	1
C	n_3–n_4	2
D	n_5–n_6	3
E	n_7–n_8	4
F	n_8+	5

Appendix C: Quality Checklist Scores

Downs and Black domain	Reporting										External Validity			Internal Validity: bias							Internal Validity: confounding (selection bias)						Power	
Downs and Black criteria number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	Total
Arnow et al., 2013	1	1	1	1	1	1	1	0	1	1	0	0	1	0	0	1	1	1	0	1	1	0	1	0	0	1	0	17
DeRubeis & Feeley (1990)	0	1	1	1	0	1	1	0	1	0	0	0	1	0	1	1	1	1	1	1	1	0	1	0	0	0	0	15
Feeley, DeRubeis, & Gelfand, 1999	1	1	1	1	1	1	1	0	1	0	1	0	1	0	1	1	1	1	0	1	1	0	0	0	1	1	0	18
Hardy et al., 2001	1	1	1	1	1	0	0	0	1	0	0	0	1	0	0	1	1	1	1	1	0	0	0	0	1	1	0	14
Haug et al., 2016	1	1	1	1	1	1	1	0	1	1	1	0	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	23
Huppert et al. (2013)	1	1	1	1	1	1	1	0	1	0	0	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	0	21
Klein et al. (2003)	0	1	1	1	1	1	1	0	1	1	1	0	1	0	1	1	1	1	0	1	1	0	1	0	1	1	0	19
Lorenzo-Luaces, DeRubeis & Webb (2014)	1	1	1	0	1	1	1	0	0	1	0	0	0	0	0	1	0	1	0	1	0	0	1	1	1	1	0	14
Marmar, Gaston, Gallagher, & Thompson (1989)	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	1	1	1	1	1	1	0	1	0	0	0	0	14
Muran et al. (1995)	1	1	1	1	1	1	1	0	0	0	1	0	1	0	1	1	0	1	0	1	1	0	0	0	0	1	0	15
Preschel, Maercker & Wagner, (2011)	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0	1	1	1	0	1	1	1	1	0	1	1	0	21
Saatsi, Hardy & Cahill (2007)	1	1	1	1	1	1	1	0	1	0	1	1	1	0	0	1	1	1	1	1	0	0	0	0	1	1	0	18
Safran & Wallner (1991)	1	1	1	1	1	1	1	0	0	0	1	0	1	0	0	1	1	1	1	1	1	0	0	0	1	0	0	16
Snippe et al., 2015	1	1	0	0	0	1	1	0	0	1	0	0	1	0	0	1	1	1	1	1	1	0	1	0	0	1	0	14
Strunk, Cooper, Ryan, DeRubeis	1	1	1	1	1	1	0	0	0	1	0	0	1	0	0	1	1	1	1	1	1	0	0	1	1	1	0	17

& Hollon (2012)

Tang & DeRubeis (1999)	1	1	0	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	6
Trepka, Rees, Shapiro, Hardy & Barkham (2004)	1	1	1	1	1	1	1	0	0	1	1	0	1	0	0	1	0	1	1	1	1	0	0	0	1	1	0	17
Downs and Black domain	Reporting										External Validity			Internal Validity: bias							Internal Validity: confounding (selection bias)						Power	
Downs and Black criteria number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	Total
Weck, Grikscheit, Jakob, Hofling & Stangier (2015)	1	1	1	1	0	1	1	0	0	1	0	0	1	0	1	1	0	1	0	1	0	0	1	0	0	0	0	13
Weiss, Kivity & Huppert (2014)	1	1	1	1	1	1	1	0	1	1	0	0	1	0	1	1	0	1	1	1	1	0	0	0	0	0	0	16
Zuroff & Blatt (2006)	1	1	1	1	1	1	1	0	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	0	23

Appendix D: Quality checklist scores double rating

Author/Rater	First Rater	Second Rater
Arnow et al., 2013	17	16
Feeley, DeRubeis, & Gelfand, 1999	18	16
Lorenzo-Luaces, DeRubeis & Webb (2014)	14	16
Strunk, Cooper, Ryan, DeRubeis & Hollon (2012)	17	18
Zuroff & Blatt (2006)	23	23

This page is left intentionally blank

Part Two: Research Report

Working alliance in Cognitive Behaviour Therapy and Counselling for Depression: A comparison of therapies and relationship to outcome

This page is left intentionally blank

Abstract

Objectives. Though the relationship between Working Alliance (WA) and outcome of psychotherapy is well established, there are aspects of the relationship that are not yet fully understood. This includes the importance of WA for different therapies, the underlying mechanisms of alliance and the impact of client factors on the alliance-outcome relationship. This research aimed to compare the Working Alliance for two therapies (Cognitive Behaviour Therapy [CBT] and Counselling for Depression [CfD]) within primary care services. The relationship of alliance and other clinical and demographic variables to outcome for the full sample was also explored. An exploratory investigation of the process of coding alliance from the perspective of observers was also undertaken.

Methods. A subset of data was taken from a cohort of participants who had taken part in the PRaCTICED trial. Twenty participants who had completed CfD and twenty participants who had completed CBT who met the inclusion criteria were selected for the study. WA for sessions one, three and five of therapy was coded by trained observers from audio recordings of therapy using the Working Alliance Inventory – Observer form (WAI-O). Demographic and outcome data (PHQ-9) was collected as part of the wider trial. All coders were asked to complete a questionnaire to capture their experiences of rating therapy using the WAI-O.

Results. Overall WAI-O and the subscales of Goal and Task were higher for CBT than CfD with medium effect sizes, though not all results achieved statistical significance. For the full sample and CfD subsample, only first session PHQ-9 had a relationship (medium effect size) to the final session PHQ-9 score and all other variables had small effect sizes. No multiple regression was therefore undertaken. For CBT, there was a significant relationship between number of sessions completed and final session PHQ-9, with a medium effect size. Though non-significant, WA and first session PHQ-9

achieved medium effect sizes. An exploratory multiple regression was undertaken with the three variables that achieved medium effect sizes. The resulting model explained 52% of the variance in final session PHQ-9 scores $R^2 = 0.52$, $F(3, 19) = 5.71$, $p = 0.01$.

The coder questionnaires highlighted some positive aspects of coding the alliance, including the measure being easy to use and the instruction manual being useful. They also highlighted several issues with coding the alliance, including difficulty capturing tensions within therapy, difficulty scoring session one of CBT and difficulty in scoring the Goal and Task subscales for CfD. These have the potential to be confounding variables when scoring WA.

Conclusion. The study provides some support for the importance of WA for the outcome of CBT. Though the results for CfD showed lower alliance scores compared to CBT and no significant relationship with outcome, there are several potential confounding variables that may have impacted on the results.

Practitioner Points

- Though the results about CfD were inconclusive, practitioners would benefit from routinely utilising measures of WA in their therapeutic work. This might help to identify any difficulties with alliance that might be addressed in therapy.
- When selecting alliance measures, practitioners would be well advised to consider the theoretical orientation of the measure chosen, to ensure that it is in keeping with the therapeutic approach undertaken.
- Further research is needed to understand the importance of alliance for CfD using larger sample sizes. In addition, exploration of the appropriateness of the WAI-O for different therapies (and whether this can be remedied via further training) would further assist our understanding of the alliance concept.

Limitations

- The sample size employed in the research was small and underpowered for both

hypothesis one and two which limits the findings significantly.

- The coder questionnaires highlighted several issues with coding using the WAI-O which may have confounded alliance ratings. It was not possible to determine whether the findings for hypothesis one were due to true differences in the strength of the alliance between the two therapies, the measure itself or lack of training and experience on the part of the coders.

The Working Alliance

The working alliance (WA) is one of the most commonly investigated topics in psychotherapy research (Doran, 2014). Whilst ways of conceptualising the WA have differed (Elvins & Green, 2008), WA has generally been described as representing the collaborative parts of the therapeutic relationship, thus involving both therapist and client (Constantino, Castonguay & Schut, 2002). Perhaps the most influential writer in the development of the WA concept was Bordin (1979) who developed a pantheoretical definition of WA consisting of the therapist and client agreeing upon therapy goals (goal), negotiating how to achieve those goals (task), as well as the bond between the therapist and client (bond).

The relationship of WA to outcome in psychotherapy has been investigated across therapy types, presenting problems and client groups (Kadzin, Marciano & Whitley, 2005; Horvath & Symonds, 1991). More recently, investigation has focussed on the WA when psychotherapy is delivered via telephone, online or video conferencing (Preschel, Maercker & Wagner, 2011). Meta-analyses of individual psychotherapy have consistently found a modest relationship between alliance and outcome across therapy types, with effect sizes ranging from .22 to .29 (Flückiger, Del Re, Wampold, Symonds, & Horvath, 2012; Horvath & Bedi, 2002; Horvath & Symonds, 1991). Though this does not reflect a strong correlational link, as Flückiger et

al. (2012) point out, this represents a stronger relationship than found for other variables, such as therapist competence or adherence to therapeutic techniques (see Webb, DeRubeis & Barber, 2010). These findings have led researchers to suggest that the alliance is essential to therapeutic change (Horvath & Bedi, 2002).

In spite of the large body of research investigating alliance, there are a number of issues with the measurement, research methodology and complexity of the alliance concept that are yet to be fully addressed by research.

Measurement of alliance. The theoretical diversity in conceptualisations of the alliance is reflected in the "proliferation" of measures developed to assess it (Elvins & Green, 2008). What they have in common is their measurement of client-therapist collaboration (Horvath, Del Re, Flückiger & Symonds, 2011).

Measures of alliance have been criticised for their focus on collaboration, thus ignoring important negative feelings in therapy (Doran, Safran, Waizmann, Bolger & Muran, 2012). This obscures important psychotherapeutic processes such as negotiation (Doran, 2014) and the repair of ruptures, which can strengthen the WA (Safran & Muran, 1990). Furthermore, it has been suggested that some alliance measures are less pantheoretical than proposed. For example, Webb, DeRubeis, Amsterdam, Shelton and Hollon (2011) have suggested that the Working Alliance Inventory (WAI) subscales of goal and task might be more suited to cognitive behavioural therapies where goals and tasks are explicitly explored. There is a need, therefore to further understand how appropriate measures are for different therapies.

The most commonly used measures have been developed to have therapist, client and observer versions (Elvins & Green, 2008). Each of these approaches has their individual strengths and weaknesses. Therapist rated alliance has been shown to be the least predictive of outcome (Horvath & Symonds 1991) often showing little correlation with client rated alliance (e.g. Marmar, Gaston, Gallagher & Thompson, 1989). Client

rated measures of alliance, whilst being more predictive of outcome, are often used alongside client rated measures of outcome. It has been argued that this creates shared variance, thus increasing the potential for type one errors (Shirk & Karver, 2003). The use of observers to rate alliance overcomes these issues. In addition, the use of these measures offers an opportunity to investigate the process of rating alliance from the perspective of an impartial observer. This has the potential to investigate issues such as how appropriate measures are for different therapies. However, there is a need to examine the therapeutic allegiance of coders, which can influence alliance ratings (Raue, Putterman, Goldfried, & Wolitzky, 1995).

Research methods. Research methods used to examine the alliance-outcome relationship have come under scrutiny in recent years (Doran, 2014). Single session ratings of alliance, particularly from session three of therapy have frequently been used (Horvath & Symonds, 1991). This has been criticised for being an arbitrary method (Kivlighan & Shaughnessy, 1995). Using generalisability coefficients, Crits-Christoph, Gibbons, Hamilton, Ring-Kurtz & Gallop (2011) assessed the impact of aggregating alliance measurements and concluded that while single session alliance measurement accounts for approximately 5% of the outcome variance, a minimum of two alliance measurements was required to achieve acceptable generalisability coefficients, with a minimum of four alliance measurements being optimal. The authors found evidence of reverse causality later in treatment (where WA is predicted by prior symptom improvement rather than vice versa), and suggest that earlier measurements should be used to predict outcome (Crits-Christoph et al., 2011).

Complexity of the alliance. Despite the authors of meta-analyses declaring the alliance-outcome correlation "ubiquitous" (Flückiger et al., 2012), the aggregation of studies using different research methodologies may obscure complex factors that influence the alliance-outcome relationship (Siev, Huppert & Chambless, 2009). It has

been suggested that the alliance-outcome correlation may be due to client contributions, therapist contributions, match between client and therapist or improvement prior to the alliance assessment (DeRubeis, Brotman & Gibbons, 2005). It is unclear, for example how client factors may influence and interact with the alliance-outcome relationship. A number of client factors have been found to influence the relationship, such as age, ethnicity, social functioning, trauma history and number of episodes of mental health difficulties (Arnow et al., 2013; Klein et al., 2003, Lorenzo-Luaces, DeRubeis & Webb, 2014).

Despite the findings of Flückiger et al. (2012) that Cognitive Behaviour Therapy (CBT) did not moderate the alliance-outcome relationship, there remains controversy about the importance of alliance for cognitive therapies that do not view alliance as a vehicle for change (Siev, Huppert & Chambless, 2009). In addition, it is unclear whether the underlying mechanisms of change are different for cognitive therapies than for other therapies (De Rubeis, Brotman, & Gibbons, 2005). It has been suggested that aspects of the WA such as task and goal, may be more important for cognitive than other therapies as this is the mechanism through which therapeutic change is achieved. (Arnow et al., 2013; Webb et al., 2011). By contrast, in experiential therapies, change is thought to be achieved via the therapeutic bond (Gelso & Hayes, 1998). Therefore, it might be predicted that when rating alliance, scores for alliance subscales that tap aspects of goals and tasks might be higher for CBT than experiential therapies while ratings for bond subscales might be higher for experiential therapies. However, there is a lack of research comparing CBT to counselling or experiential therapies through which to test this hypothesis. In addition, there are inconsistent findings where alliance-outcome relationships have been compared between CBT and other therapies. For example, Snippe et al. (2015) found a significant alliance-outcome correlation for CBT but not Mindfulness based cognitive therapy (MBCT). While Raue, Goldfried, &

Barkham (1997) found that alliance ratings were higher for CBT than psychodynamic or interpersonal therapy, Klug, Zimmerman & Huber (2014) found no significant differences in client rated alliance between CBT, Psychodynamic and Psychoanalytic therapy. This suggests that further research comparing therapies, the difference in alliance subscales and mechanisms of change is needed.

Improving Access to Psychological Therapies (IAPT)

There are increasing opportunities to exploring process issues in cognitive and experiential therapies within the Improving Access to Psychological Therapies (IAPT) initiative in the UK. Since the publication of the Layard report (2006), it is recommended that commonly occurring mental health difficulties such as anxiety and depression are treated within IAPT services (NICE 2009; 2011). Though initially, IAPT services were driven by the evidence base for CBT, experiential therapies, such as Counselling for Depression are now being trialled (e.g. Saxon et al., 2017). Although WA has an impact on the effectiveness of psychotherapy interventions, there is a lack of research into WA in primary care settings in the UK. In addition, there is a need to examine the impact of WA and outcome for newly developed therapies. This research therefore aims to investigate the WA in CfD and CBT at a number of time points and investigate their relationship both to one another and to treatment outcome.

Aims

1. To determine if there is a difference in WA ratings between CfD and CBT.
2. To determine to what extent WA accounts for outcome when other variables are accounted for (gender, employment status, age, clinical risk, number of sessions completed).
3. To complete an exploratory examination of the process of coding alliance from the perspective of observers.

Hypotheses

1. WA ratings will be higher for CBT than CfD
    a. The task subscale of the WAI will be higher for CBT than CfD
    b. The goal subscale of the WAI will be higher for CBT than CfD
    c. The bond subscale of the WAI will be higher for CfD than CBT
2. WA will be a significant predictor of outcome after controlling for other variables (gender, employment status, age, clinical risk, number of sessions completed).

Method

Ethics and Data Protection

Ethical approval was granted to the PRaCTICED trial on 27.04.14 (REC ID: 14/YH/0001; IRAS ID: 130352). A copy of the approval letter is in Appendix E. During the recruitment process for the trial, clients were given information (Appendix F) that included information about the process research that recordings might be used for. All clients completed consent forms to be recorded and for their data to be used for research purposes (Appendix G)

The PRaCTICED Trial

Data was taken from participants already recruited to the PRaCTICED trial (Saxon et al., 2017). The PRaCTICED trial is a pragmatic non-inferiority randomised trial of the clinical and cost effectiveness of counselling for depression (CfD) versus cognitive behaviour therapy (CBT), for clients in primary care meeting a diagnosis of moderate or severe depression. The trial is being conducted in the Sheffield Improving Access to Psychological Therapies (IAPT) service and aims to treat 500 people presenting with moderate or severe depression. Participants were required to meet an

ICD-10 diagnosis of moderate or severe depression using the Clinical Interview Schedule-Revised (CIS-R; Lewis, 1994), carried out by an independent assessor. Exclusion criteria included presence of prior diagnosis of personality disorder, bipolar disorder or schizophrenia.

Participants

Forty participants were selected: twenty from each of the CBT and CfD arms of the PRaCTICED trial by a researcher on the project who was blind to therapy outcome. The inclusion criterion was that therapy recordings had to be available for sessions one, three and five. All participants, therefore had received a minimum of five sessions of therapy. This was to enable the research to examine the progress of WA in the early to mid-stages of therapy and to ensure that participants had received a 'dose of therapy' that might be sufficient to make symptomatic change (Kadera, Lambert & Andrews, 1996). Of those participants meeting the inclusion criterion, they were selected so as to use as wide a range of therapists as possible. This was in order to minimise the impact of therapist effects on outcome which have been observed to impact the alliance-outcome correlation (Del Re, Flückiger, Horvath, Symonds & Wampold, 2012).

Treatments

CfD: CfD is a therapy designed for use within IAPT services for clients experiencing depression. It is delivered by counsellors and aims to incorporate person-centred and emotion-focussed approaches (Saxon et al., 2017). The curriculum was developed by the British Association for Counselling and Psychotherapy and follows the text by Sanders & Hill (2014). Prior to the trial CfD training was provided to all counsellors. This comprised a five day taught component with 80 hours supervised practice.

CBT: CBT within high intensity IAPT services comprises two protocol driven interventions: Beckian Cognitive Therapy and Martell's behavioural activation. It is

delivered by high intensity CBT practitioners. CBT therapists were trained in Beckian CBT when they completed their IAPT training. In addition, regular top-up workshops were provided. The delivery of CBT is in accordance with Cognitive behaviour therapy: Basics and beyond (Beck, 1995).

Treatments of up to 20 sessions were administered for both therapies as part of the PRaCTICED trial. Treatment fidelity was ensured according to the PRaCTICED trial protocol (Saxon et al., 2017). This involved the calibration of competence raters against a national expert. Raters completed competence ratings of randomly selected therapy tapes according to a sampling strategy designed to ensure that a representative number of tapes from each therapist, at different stages of therapy were included. The treatment fidelity results are not available as the trial is still in progress.

Measures

Working Alliance. WA was measured using the Working Alliance Inventory Observer Form (WAI-O). The Working Alliance Inventory is a pantheoretical measure developed by Horvath and Greenberg (1986). It is a 36-item measure with items rated on a seven-point Likert scale, yielding a maximum total score of 252. The WAI yields three subscales: goal, task and bond, based on Bordin's (1979) theory of alliance. The observer version of the WAI was developed by Tichenor and Hill and requires no training for coders (1989; Appendix I). The WAI-O has been demonstrated to have good internal consistency with coefficient alphas reported of 0.97 and 0.96 respectively (Hanson, Curry & Bandalos, 2002; Tichenor & Hill, 1989). It also has good interrater reliability, with intraclass correlations of 0.79 and 0.92 reported respectively (Hanson et al., 2002; Tichenor & Hill, 1989). Permission has been given for the use of the WAI-O in this study (Appendix J). The observer form was chosen because of methodological difficulties with the use of both client and therapist ratings of WA: therapist ratings of WA have been found to be less reliable measures of WA (Horvath & Symonds 1991)

and there are issues with outcome and alliance being measured by the same person as this creates shared variance, and increases the risk of type I errors (Elvins & Green, 2008).

The instructions for the WAI-O were developed by Raue, Castonguay & Goldfried (1993) and provided by Professor Raue in response to a direct request. Though a more recent manual has been developed (Darchuk et al., 2000), the manual changes both the wording of the Likert scales and the anchor point of ratings. No research could be found validating this version of the WAI-O and there was no response to requests to the authors for validation data. In addition, no studies could be found that had used this version of the measure.

Outcome. Therapy outcome was measured using the Patient Health Questionnaire, abbreviated to PHQ-9 (Kroenke, Spitzer & Williams, 2001; Appendix K). The PHQ-9 was completed by each client recruited to the PRaCTICED trial at intake and on a sessional basis. The PHQ-9 completed at the final session was utilised as the measure of overall outcome.

The PHQ-9 is a nine-item measure of the symptoms of depression that a client has experienced over the preceding two weeks. Each item is rated in terms of symptom frequency from 0 = not at all to 3 = every day yielding a maximum score of 27 indicating greatest difficulty. A score of ten indicates moderate depression, with higher scores indicating greater severity (Kroenke, Spitzer & Williams, 2001).

The PHQ-9 has been validated on a large sample and has been found to have good test-retest reliability (0.84) and good criterion and construct validity (Kroenke, Spitzer & Williams, 2001). It has been found to have good associations (r = .73) with depression severity on other measures such as the BDI (Martin, Rief, Klaiberg & Braehler, 2006). It also has the benefit of being quick to use and does not require training for the person administering it.

Additional Data. Clinical Outcomes in Routine Evaluation – Outcome Measure (CORE-OM; Evans et al., 2000; Appendix L). The CORE-OM is a 34 item measure of psychological distress for use across presenting problems and therapeutic approaches. The measure yields four subscales: wellbeing, symptoms, functioning and risk. The CORE-OM was completed by all participants at intake. The CORE-OM has good internal reliability (between 0.75 and 0.95), good test-retest stability (0.87-0.91) and has shown large differences between clinical and non-clinical samples. (Evans et al., 2002). For the purposes of this study, the CORE risk score was examined for hypothesis two during the pre-regression stage.

The extent to which additional client factors which may account for outcome other than WA were investigated for hypothesis two. This data was collected routinely as part of the PRaCTICED trial, for example employment status, gender, age, ethnicity and number of sessions completed. These were examined for hypothesis two during the pre-regression stage.

To capture aspects of the process of coding to inform future research and training, coders using the WAI-O were asked to complete a feedback questionnaire (Appendix M) within one month of completion of the project. It was hoped that this would highlight any issues with the training and coding process that could be addressed in future research as well as elucidating aspects of the research findings.

Power Analysis

Effect sizes for the comparison of alliance between two therapies have ranged from small (e.g. Snippe et al., 2015) to large (e.g. Raue, Goldfried & Barkham, 1997). For hypothesis one, assuming *p* values of .05, when completing an independent samples analysis for two groups, when predicting a medium effect size, a sample size of n = 64 per group is recommended (Cohen, 1992). For hypothesis two, a linear regression was undertaken. Though there is disagreement in the literature about the number of Subjects

Per Variable (SPV) required for a linear regression, it has been suggested that a minimum SPV of between two and five is sufficient (Austin & Steyerberg, 2015; Green, 1991). A sample size of 40 will therefore be sufficient to include up to eight variables in the regression model (using the more conservative SPV).

Procedure

Working alliance ratings were coded for sessions one, three and five of therapy: 120 therapy sessions in total. The coding of the WAI-O was completed by four post-graduate psychology students recruited from Sheffield University. The interview process included a task designed to assess coders' ability to identify important factors in the WA.

The initial training for coders was an eight-hour course, facilitated by the author (Appendix N). It consisted of the background to the research study, the background to the WAI-O and how to use it, data protection issues and the importance of self-care. A practical exercise involving using the WAI-O to rate twenty-minute extracts from therapy sessions was used until calibration was achieved. Discrepancies of more than one point on the seven-point Likert scale were discussed as a group following the method used by Raue, Castonguay and Goldfried (1993). Sufficient calibration was achieved following the rating of two twenty-minute extracts. In addition, all coders were given the coding manual (Raue, Castonguay & Goldfried, 1993) and signed a confidentiality form (Appendix O). All coders were required to complete online information governance training provided by School of Health and Related Research (SCHaRR).

Twenty percent of therapy recordings were scored by two coders: twenty-four in total. This was to maintain reliability of scoring and prevent drift. Coders were encouraged to make note of their discussions on a form (Appendix P). Double rated recordings were allocated so that each coder double coded a therapy recording for at

least every five recordings completed. Each coder double rated against one another four times in total. This data was used to calculate inter-rater reliability (see results). For double coded recordings, the average of the two WAI-O scores was used for the analysis. In addition, monthly meetings were held to discuss any issues with scoring, during which a further 20 minute excerpt from a therapy recording was coded by the group and discussed.

Including double rated recordings (see below), 144 hours of therapy recordings were coded. One coder completed slightly more coding hours than the other three due to being recruited onto a scheme which gave them 100 hours to work on the project (n = 54) vs (n=30). Each coder was allocated to an equal number of CBT and CfD clients. Therapy tapes were allocated to the coders at random using research randomiser (Urbaniak & Plous, 2013). All coders and the author were blind to outcome information and participant identity during the coding process. In addition, the coders were blind as to whether they were listening to CBT or CfD (though they may have learned the different styles of the therapy). Coders were instructed to stop listening to the therapy session if they knew either the client or therapist. This happened on one occasion as the coder knew the therapist. The coder stopped listening immediately, and the recording was assigned to another coder.

All client information was kept in a locked filing cabinet at the School of Health and Related Research (SCHaRR) where the coding was completed. Audio recordings of therapy were downloaded onto encrypted memory sticks individually by SCHaRR staff and deleted after each coding was completed. Coders used headphones in a private office. Each participant and therapist was identified via a unique number and not their name on both therapy recordings and completed WAI measures.

Data Analysis

Tests of normality were carried out on the overall mean WAI scores, subscale

scores, intake age of participants, intake risk scores, number of sessions completed and PHQ-9 scores. Shapiro-Wilk's tests (p>0.5) (Shapiro & Wilk, 1965; Razali & Wah, 2011) and visual inspection of histograms, normal Q-Q plots and box plots suggested that overall WAI scores, subscale scores and of Goal, Task and Bond and additional clinical variables were all non-normally distributed. PHQ-9 scores were approximately normally distributed. Visual inspection of histograms suggested that the WAI overall score and subscale scores were skewed towards higher scores, suggesting ceiling effects in the measure. Visual inspection of histograms for CfD and CBT conditions separately suggested that there were differential patterns of distribution of alliance scores between the two therapies with a greater spread of scores for CfD. To compare the demographic and clinical variables for the two therapies, Chi Squared, Mann-Whitney U tests and independent samples T-Tests were undertaken as appropriate. In addition, a Friedman test of repeated measures was performed on the sessional WAI data to examine any differences between sessions one, three and five.

For hypothesis one, Mann-Whitney U tests were conducted on the overall mean WAI score and subscale scores for CfD and CBT to compare the differences. The effect size (r) was calculated for each analysis using the formula $r = Z/\sqrt{N}$.

In addition, to account for the differential spread of scores between the two therapies, the percentage of scores above and below the overall mean score for CfD and CBT were calculated and the difference between the two therapies examined using a Chi Square test. This analysis was performed for the overall WAI scores and the three subscale scores.

For hypothesis two, a multiple linear regression analysis was undertaken. The first step was to complete correlation analyses for all demographic and clinical variables against the outcome variable of final PHQ-9 score. Note that for the PHQ-9, a lower score indicates lower symptom severity. First session PHQ-9 was included in order to

control for symptom severity at the start of therapy. As none of the variables for the full sample were significantly correlated with the outcome variable, the correlation analyses were repeated for the individual therapies separately. As there was only a medium effect size for one of the variables (first session PHQ-9) and none of the other variables for CfD, a regression analysis was not completed for the CfD sample. For the CBT sample, an exploratory multiple linear regression was undertaken for the three variables that showed a medium effect size. The choice of effect size as opposed to a statistically significant *p* value was suboptimal though it accounted for the small sample size. For the final model, analyses of the assumptions of normality, linearity multicollinearity and homoscedasticity were undertaken.

Results

Inter-rater reliability

Inter-rater reliability was calculated for each pair of coders using a one-way random effects, absolute, single measures intra-class correlation (ICC: McGraw & Wong, 1996) to assess the degree to which each pair of coders were consistent in their WAI ratings. One outlier pair of scores between coder 1&3 was removed from the analysis. The resulting ICC's were then averaged following the procedure used by Krupnick et al. (1996). The resulting score suggested good interrater reliability, ICC = 0.74 (Cicchetti, 1994). The individual scores for each coder pair is shown in Table 1. Interrater reliability between coder pairs ranged from fair to excellent (Cicchetti, 1994).

Table 1.
Inter-rater reliability between coders

Coder pair	ICC	Confidence intervals
1&2	0.96	.65 - 1.00
1&3	0.45	-.72 --0.98
1&4	0.77	-.12 - 0.98
2&3	0.94	.49 - 1.00
2&4	0.55	-.49 - 0.96
3&4	0.79	-.80 - 0.99

Sample Characteristics

Characteristics of the sample are shown in table 1. The sample consisted of 40 adult participants aged between 19 and 66. The CfD subgroup included 9 women and 11 men aged between 20 and 66 (mean = 41; *SD*= 12) and the CBT subgroup included 12 women and 8 men aged between 19 and 55 (mean = 43; *SD* = 10). The mean number of sessions completed was 12.9 for CfD and 12.75 for CBT. The majority of clients in the sample completed therapy (*N* = 37). The majority of the sample were White British (*N* = 34) and employed (*N* = 23). Data on marital status was not collected as part of the trial.

Table 2.
Participant demographic and clinical variables

	CfD (*N* =20)		CBT (*N* = 20)	
Variable	*N*	*%*	*N*	*%*
Gender				
Female	9	45	12	60
Male	11	55	8	40
Ethnicity				
White British	17	85	17	85
White - other	1	5	0	0
Afro-Caribbean	0	0	1	5
Mixed - White & Afro-Caribbean	2	10	1	5
Not disclosed	0	0	1	5
Employment				
Employed	12	60	11	55
Sick leave	3	15	2	10
Unemployed	3	15	4	20

Student	1	5	0	0
Homemaker	0	0	3	15
Retired	1	5	0	0
Therapy completion status				
Completer	18	90	19	95
Dropout	2	10	1	5
	M	*SD*	*M*	*SD*
Age	41.7	11.85	42.7	10.91
Intake risk score (CORE risk)	4.25	5.5	5.58	5.78
Intake severity score (PHQ-9)	18	4.15	18.75	5.38
Number of sessions completed	12.9	4.23	12.75	4.18

A Chi-Squared test of independence suggested no significant difference in gender between the two groups (X^2 (1) = 0.902, p = .342). In terms of ethnicity, employment and completer status, there were insufficient data within each category to complete a Chi-Squared analysis. However, it is clear that there is very little difference in the ethnicity, employment status or completer status between the two groups.

There was no significant difference in the age of participants between therapies (U = 182, p = .626). There was no significant difference in intake PHQ-9 scores between CfD and CBT t(38) = -0.49, p = .63. There was no significant difference in intake risk (CORE risk subscale) score between the two groups (U = 169.5, p = .401). There was no significant difference in the number of sessions completed between therapies (U = 194, p = .871).

Therapist Characteristics

CfD was provided by 11 therapists, with a range of one to six clients from the sample (mean = 1.82; SD = 1.54). Therapist information was not available for one CBT client. For the remaining participants, CBT was provided by 12 therapists, with a range of one to four clients each from the sample (mean = 1.58; SD = 0.90). It was not possible to complete an investigation of therapist effects due to the small numbers of participants per therapist. The results of therapist competence ratings were not available due to the PRaCTICED trial still being in progress.

Alliance over the course of therapy

Table two shows the WAI overall scores over the course of therapy for the two therapies. A non-parametric Friedman test of repeated measures was conducted to examine whether there were significant differences in overall WAI scores at sessions 1, 3 and 5. The test was non-significant for CfD X^2 (2, 20) = 1.9, p = .387, for CBT X^2 (2, 20) = 0.514, p = .774 and for the full sample combined X^2 (2, 20) = 0.948, p = .622.

Table 3.
Working alliance scores over the course of therapy

	CfD (*N*= 20)				CBT (*N*=20)			
Variable	Mean	SD	Median	Range	Mean	SD	Median	Range
WAI total score	222	16.29	225.0	191-246	232	10.9	234.0	210-247
Session 1	227	14.16	228.0	194-248	233	14.34	236.0	206-252
Session 3	217	23.00	217.5	168-246	234	13.24	238.5	210-250
Session 5	223	22.00	230.0	175-250	232	10.9	236.5	210-247
WAI subscale: goal								
Session 1	74	5.79	75.0	63-83	78	4.78	80.0	68-84
Session 3	70	8.90	72.0	53-82	77	5.28	79.0	68-84
Session 5	73	8.20	76.0	60-84	76	6.98	78.0	57-84
WAI subscale: task								
Session 1	74	5.30	75.0	67-82	77	5.53	78.0	64-84
Session 3	71	8.67	72.0	54-82	77	5.61	78.0	67-84
Session 5	73	7.65	75.0	58-83	76	6.64	78.0	68-83
WAI subscale: bond								
Session 1	78	4.76	80.0	64-83	78	5.68	80.5	64-84
Session 3	76	6.47	76.5	59-84	79	3.74	79.5	73-84
Session 5	77	7.67	78.5	59-84	79	7.33	81.0	66-85

Outcome

A paired samples t-test indicated that there was a significant reduction in PHQ-9 scores from between the first and last session of therapy $t(39) = 7.10$, $p<0.01$. There was also a significant reduction in PHQ-9 scores for CfD $t(19) = 5.25$, $p<0.01$ and CBT $t(19) = 4.69$, $p<0.01$ individually. The means and standard deviations for each therapy and the overall sample are in table four.

The cut off for determining clinically significant change or "caseness" on the PHQ-9 is a score ≤10 (Kroenke, Spitzer & Williams, 2001). By this determinant, 50% of

clients made clinically significant change (nine for CfD and eleven for CBT).

The reliable change index (RCI) was calculated using normative data published by Kroenke, Spitzer and Williams (2001). Overall, 13 clients showed no change, one deteriorated and 26 showed reliable improvement. For CfD, six clients showed no change and fourteen showed reliable improvement. For CBT, seven clients showed no change, one client showed deterioration and twelve showed reliable improvement.

Table 4.
PHQ-9 scores at session one and final session

	Session one PHQ-9			Final PHQ-9		
	Mean	SD	Range	Mean	SD	Range
CFD	18.05	4.15	10-24	10.40	6.58	0-23
CBT	17.45	4.43	8-25	10.05	6.99	2-27
Overall	17.75	4.25	8-25	10.23	6.66	0-27

Hypothesis One

An independent samples Mann-Whitney test showed no difference in intake severity between CfD and CBT, therefore it was not necessary to control for intake severity in the analysis for hypothesis one. In addition, a Chi-Squared test of independence indicated that there was no significant difference between the frequency of PHQ-9 severity categories at intake (none, mild, moderate, moderately severe, severe) between CfD and CBT (X^2 (1) = 0.642, p = .887).

A Mann-Whitney U test indicated that mean WAI scores were non-significantly higher for CBT than CfD (U = 129, p = .056) with a medium effect size. The Goal subscale was significantly higher for CBT than CfD (U = 118, p = .026) with a medium effect size, as was the Task Subscale (U = 119, p = .028). The Bond subscale was non-significantly higher for CBT than CfD (U = 164.5, p = .341) with a small effect size. The means, standard deviations, medians, ranges and effect sizes are shown in table 3. Medium effect sizes were found for all of the variables with the exception of the Bond subscale.

Table 5.
Averages, SD's, medians, ranges and effect sizes for average working alliance scores

	CfD				CBT				ES(*r*)
	Mean	SD	Median	Range	Mean	SD	Median	Range	
Overall	22	16.29	223	191-246	232	10.90	236	210-247	.31
Goal	73	8.20	74	62-82	77	6.99	78	68-82	.36
Task	73	7.65	73	62-81	77	6.64	77	68-83	.35
Bond	77	6.47	78	65-83	79	7.73	79	71-83	.15

For the second analysis, a Chi-Squared test of independence was undertaken comparing the frequency of WAI total and subscale scores above and below the average score for CfD and CBT combined. The results are in table 3. A significant relationship was found for the total WAI score (X^2 (1) = 5.01, p = .025) suggesting that significantly more clients undertaking CBT had above average WAI scores. The effect size was medium (phi = .35; Cohen, 1988). For the goal subscale, though there were more above average scores for CBT than CfD, this relationship was non-significant (X^2 (1) =3.636, p = .057); however the effect size was medium (phi = .3; Cohen 1988). For the task subscale, a significant relationship was found (X^2 (1) = 5.013, p = .025), suggesting that significantly more clients undertaking CBT had above average goal scores. The effect size was medium (phi = .35; Cohen, 1988). For the Bond subscale, there was no significant difference between the two therapies (X^2 (1) = .921, p = .337) with a small effect size (phi = .152; Cohen, 1988).

Table 6.
Frequency of working alliance scores above and below the average score for CfD and CBT

	CfD		CBT		ES (phi)	*p*
	Above	Below	Above	Below		
Overall WAI	8	12	15	5	0.354	0.025*
Goal	8	12	14	6	0.302	0.057
Task	8	12	15	5	0.354	0.025*
Bond	10	10	13	7	0.152	0.337

Note. * $p \leq 0.05$

Hypothesis Two

Table seven shows the correlations and *p* values for the demographic and clinical variables with the outcome variable (final session PHQ-9). For the variable "ethnicity" the categories were reduced to "White British" or "non-White British" and the variable "employment" was reduced to "currently working" or "not currently working". As there was no significant difference in session one, three and five WAI scores, the mean WAI score over the three sessions was entered as a variable. As table seven shows, when correlation analyses were completed for all of the demographic and clinical variables, only first session PHQ-9 was significant at the $p \leq 0.05$ level. Therefore, no multiple regression analysis was completed for the full sample.

Table 7.
Correlations of variables for full sample with final session PHQ-9

Variable	*r*	*p*
Working alliance	-.07	0.68
Number of sessions	-.14	0.38
Therapy type	-.27	0.87
First session PHQ-9	-.30	*0.05
Intake risk score	.21	0.12
Age	.04	0.80
Ethnicity	.05	0.77
Employment	.08	0.63
Gender	-.02	0.90

Note. * $p \leq 0.05$

Although therapy type was not significantly correlated with outcome, the pattern of difference in WAI scores was sufficient to justify investigating the correlations of demographic and clinical variables for the two therapies separately.

Table 8.
Correlations of variables for CfD and CBT with final session PHQ-9

	CfD		CBT	
Variable	*r*	*p*	*r*	*p*
Working alliance	.14	0.53	-.36	0.12
Number of sessions	.17	0.47	-.45	0.05*
First session PHQ-9	.32	0.17	.30	0.21
Intake risk score	.19	0.41	.24	0.30

Table 8.
Correlations of variables for CfD and CBT with final session PHQ-9 continued

	CfD		CBT	
Variable	*r*	*p*	*r*	*p*
Age	-.02	0.92	.11	0.64
Ethnicity	.05	0.84	.04	0.85
Employment	-.13	0.58	.28	0.24
Gender	.05	0.82	-.08	0.72

Note. * $p \leq 0.05$

As table eight shows, for CfD, only first session PHQ-9 achieved a medium effect size (Cohen, 1988) and none of the variables achieved statistical significance at the level of $p \leq 0.05$ and all other effect sizes were small. A regression for the CfD subgroup was not therefore completed. For CBT, three variables achieved a medium effect size (working alliance, number of sessions completed and first session PHQ-9). However, only number of sessions completed achieved statistical significance at a level of $p \leq 0.05$.

An exploratory multiple regression was completed for the CBT data. As number of sessions completed had the strongest relationship to outcome, this was entered at step one. At step two, WAI score and first session PHQ-9 score were both entered, since they had similar effect sizes. The resulting model is shown in table nine.

The results were examined for the assumptions of no multicollinearity, normality, autocorrelation, homoscedasticity and linear relationship. The results suggested that all assumptions were met, although the lack of data points made the quality of the linear relationship difficult to determine.

Table 9.
Summary of coefficients in model predicting final session PHQ-9

Step	Variable	*B*	*SE(B)*	*β*	*t*	Sig.
1	Number of sessions	-0.99	0.31	-0.59	-3.24	0.01*
2	Working alliance	-0.21	0.11	-0.33	-1.87	0.08
	First session PHQ-9	0.71	0.29	0.45	2.46	0.03*

Note. Sig. = significance; * $p \leq 0.05$

At step one, number of sessions explained 20% of the variance: $R^{2 =}$ 0.2, $F(1, 18)$ =

4.46, $p = 0.05$. The addition of working alliance and first session PHQ-9 explained an additional 32% of the variance: $R^2 = 0.32$, $F(2, 16) = 5.27$, $p = 0.02$. The final model explained 52% of the variance in final PHQ-9 scores $R^{2 =} 0.52$, F(3, 19) = 5.71, $p = 0.01$. While at step 2, the *t* value statistic was non-significant for working alliance, the same model with working alliance removed explained only 40% of the variance. Working alliance was therefore retained within the model.

Coder Questionnaires

All four coders completed a questionnaire (Appendix H) to capture feedback on the process of coding therapy recordings and to identify any issues with coding that could be addressed in further research. The findings are summarised below.

Therapy allegiance: Three out of four coders stated that they felt more positive about CBT than CfD prior to coding. The other coder had no clear allegiance. Following coding all four coders stated that they felt more positive about CBT than CfD.

Difficulties with coding (process issues). Three out of four coders said that at times they felt that clients were being compliant with therapy rather than having a positive WA with the therapist. However, they found it difficult to reflect this in the WAI-O. One coder said that in general it was difficult to reflect therapy "nuances" in the measure. In addition, one coder said that they had to give a lower score for sessions where the client and therapist disagreed, even where the resolution of the disagreement seemed to result in an improved WA.

Difficulties with coding (therapy issues). Three out of four coders said that session one of CBT was difficult to score due to the format of this session. Three out of four coders said that they felt that the measure was more appropriate for CBT than CfD. Two coders stated that they found it difficult to score Task and Goal items for CfD due to these aspects of therapy not being explicitly addressed during therapy.

Advantages and disadvantages of the measure. Coders generally felt that the measure was quick and easy to use. However, two of the coders felt that it was too long and one coder felt that the ordering of the items sometimes led to response bias.

Advice for future coders/training. All four of the coders said that the manual was useful, though one said that it could be improved by giving more detailed information about some items. Additional advice was to complete coding immediately after listening to the recoding, and to consider one's own bias when completing coding.

Discussion

This study aimed to determine whether there are differences in alliance scores between CfD and CBT and to examine whether alliance and the full sample combined were associated with outcome when other clinical and demographic variables were accounted for. In addition, an exploratory examination of the process of coding alliance by observers was completed to further understand the alliance construct and as an aid to understanding the results for hypothesis one and two. In line with hypothesis one, overall WAI scores were higher for CBT than CfD. Though the effect size for the difference between the therapies was medium (Cohen, 1988), the results were non-significant. However, significantly more WAI scores fell above the average overall score for CBT than CfD. These patterns of alliance could not be attributed to differences in ratings for sessions one, three and five as there were no significant differences between ratings for either individual therapies or the full sample combined. There was evidence of significant ceiling effects in the measure which have been previously reported (Raue, Goldfried & Barkham, 1997).

Though there is a lack of research comparing CBT to counselling or experiential therapies, these findings support several studies where alliance scores have been higher for CBT based than other therapies such as MBCT (Snippe et al., 2015);

psychodynamic or interpersonal therapy (Raue, Castonguay & Goldfried, 1993; Raue, Goldfried, & Barkham, 1997) and brief supportive psychotherapy (Arnow et al., 2013). However, these findings have not always been replicated (Klug, Zimmerman & Huber, 2014) and meta-analytic approaches have not found that treatment type moderates the alliance-outcome relationship (Horvath & Symonds, 1991; Martin, Garske & Davis, 2000).

In line with the hypotheses, the goal subscale of the WAI was significantly higher for CBT than CfD with a medium effect size. There were significantly more scores above the average overall score for CBT than CfD. The task subscale was significantly higher for CBT than CfD with a medium effect size and significantly more WAI scores were above the mean overall score for CBT than CfD. Contrary to our hypotheses, however, there were no significant differences in the bond subscale of the WAI for the two therapies and no significant difference in the number of scores that fell above the average overall score for the two therapies. These results are in line with the findings of Arnow et al. (2013) who found that the goal and task subscales of the WAI were significantly higher for CBASP (a variant of CBT) with Brief Supportive Psychotherapy but found no significant differences in the bond subscale.

Both treatments were effective at reducing symptoms with 45% of clients meeting "caseness" at the end of therapy for CfD and 55% for CBT. These results are close to the national average of 48.9% for IAPT services in 2016 (Health and Social Care Centre, 2016). Despite this, contrary to hypothesis two, correlation analyses for the full sample showed no significant relationship between WAI scores or any other variables with outcome. The effect size for alliance and outcome for the full sample was small.

These findings run contrary to studies who have found medium effect sizes in the relationship between alliance and outcome (Huppert et al., 2014; Muran et al., 1999; Weck, Grikscheit, Jakob, Hofling & Stangier, 2015) and the findings of cross therapy

meta-analyses of the alliance-outcome correlation which have found effect sizes of around .22 to .29 (Flückiger, Del Re, Wampold, Symonds, & Horvath, 2012; Horvath & Bedi, 2002; Horvath & Symonds, 1991). However, not all published research has found effect sizes of this magnitude (Snippe et al, 2015; Strunk, Cooper, Ryan, DeRubeis & Hollon, 2002) and it is possible that studies with non-significant findings remain unpublished due to the file drawer effect (Rosenthal, 1979).

When the samples were divided by therapy type, the effect sizes for CfD with outcome were small and positive (a negative correlation would be expected where higher alliance was associated with lower PHQ-9 score). The lack of findings regarding the relationship between alliance and outcome for CfD runs contrary to research that has found that alliance does influence outcome for counselling (Leibert, Smith & Agaskar, 2011) using larger sample sizes. The small sample size of the current study makes it impossible to conclude that alliance is unimportant for outcome in CfD. It is also possible that the alliance-outcome correlation was confounded by another variable, such as the choice of alliance measure; discussed below. Finally, it could be that experiential therapies are more challenging of the therapeutic relationship (as perhaps evidenced by the broader distribution of WAI scores for CfD). If this is the case, it may be that the alliance-outcome relationship is more complex for CfD than CBT.

Effect sizes for alliance and outcome for CBT were medium (Cohen, 1988). Though non-significant, effect sizes for CBT were akin to research into WA and outcome for CBT (Klein et al., 2003; Weck, Grikscheit, Jakob, Hofling & Stangier (2015) and were in line with effect sizes seen in cross therapy meta-analyses of the alliance-outcome correlation (Flückiger, Del Re, Wampold, Symonds, & Horvath, 2012; Horvath & Bedi, 2002; Horvath & Symonds, 1991). Though the lack of statistical significance may be a result of the small sample size, the findings should not be overstated.

Additional correlations showed that the only demographic and clinical variables to achieve medium effect sizes were first session PHQ-9 (added to control for intake severity) and number of sessions completed. An exploratory multiple regression with number of sessions completed, alliance and first session PHQ-9 score explained 52% of the variance in final PHQ-9 scores. This runs contrary to findings that age and ethnicity have been found to contribute significantly in models of the alliance-outcome relationship (Arnow et al., 2013; Klein et al., 2003). The other factors, however, were more exploratory in nature. The effect size found for the first session PHQ-9 with final outcome demonstrates the importance of accounting for intake severity when investigating the relationship between alliance and outcome as this accounted for a significant portion of the variance in final outcome score.

The coder questionnaires highlighted some positive aspects of coding the alliance, including the measure being easy to use and the instruction manual being useful. However, they indicated several difficulties that the coders came across whilst rating therapy recordings. Three coders reported that they were forced to rate alliance as strong, even when they felt that the client was demonstrating compliance rather than a genuine alliance. This provides support for the argument that the alliance construct currently places too strong an emphasis on agreement (Safran & Muran, 2006). In addition, one coder reported that they felt that the wording of the WAI meant that they had to score a session lower because of disagreement, even though they felt that this led to a stronger alliance. Though the process of rupture and repair is thought to be an important process of change in psychotherapy (Norcross & Wampold, 2011), this study provides limited evidence that it is difficult to reflect this in the WAI-O.

Two coders fed back that they found some aspects of CfD difficult to code: particularly the Goal and Task subscales as this is not directly addressed in the therapy. Although session one of therapy was harder to code for CBT, all of the coders felt that

the measure was more appropriate for CBT than CfD. It is unclear, however, whether this was due to a lack of understanding about CfD, which could be addressed via training. Though there is some evidence that commonly used measures such as the CALPAS, VTAS and WAI-O are broadly equivalent (Tichenor & Hill, 1989), no research could be found to date that has explored the process of coding alliance from the perspective of observers.

Theoretical implications

This study has several important theoretical implications for the measurement and conceptualisation of the alliance concept. Firstly, there is the question of whether the findings regarding alliance for CfD are due to the therapy itself, or some other factor in the measurement of the alliance. It is unclear whether all measures of the alliance are equally suitable for evaluating different therapies. Though the WAI is one of the most commonly used alliance measures (Elvins & Green, 2008) the coder feedback in this study suggest that it may be less appropriate for therapies that are more exploratory in nature and do not directly address issues of task and goals. An alternative explanation is that these aspects of alliance may be too subtle to detect in less directive therapies by coders not experienced in psychotherapy. This has the potential to be a confounding variable when examining the alliance-outcome correlation which could be addressed by providing more detailed training on different therapies.

There were significant ceiling effects in the WAI-O scores, which has been previously reported (Raue, Goldfried & Barkham, 1997). Though this has been addressed by changing the Likert scales and anchor points for scoring (by Darchuk et al., 2000), this version of the WAI-O has not been standardised and no research papers could been found that have used this version of the measure.

Clinical implications

Though there are no definitive conclusions regarding the importance of alliance

for outcome in CfD, there is some evidence that alliance is important for the outcome in CBT. Given the evidence that therapist ratings of alliance are least predictive of outcome (Horvath & Symonds 1991), it would be of benefit for therapists to use alliance measures routinely in therapy. This would allow the monitoring of alliance ruptures and address them as part of therapy. Given the tentative findings about difficulties in coding CfD using the WAI, it is important for clinicians to consider the theoretical orientation of alliance measures when choosing one to use in practice as this may impact on the results.

Strengths and Limitations

This is the first study to examine the alliance in CfD. Due to the data coming from a clinical trial, therapies were standardised according to a protocol and were monitored for therapist competence.

The study used well established measures of outcome and alliance (Elvins & Green, 2008). The alliance coders received eight hours of training and met for regular meetings to discuss issues and to complete coding as a group. A process of double coding of therapy recordings was used in order to calculate inter-rater reliability and to assist with maintaining consistency of coding. Though a method commonly used in alliance research (e.g. Weck et al., 2016; Zuroff & Blatt, 2006), the use of intra-class correlations (ICC's) for overall scores may obscure differences in scores for individual items. In addition, there was variability in ICC scores for each of the six coder pairs, ranging from 0.45-0.96. While this could be due to the small number of double rated recordings undertaken by each coder pair (four), it does cast potential doubt over the inter-rater reliability of the measure. In addition, three of the coders had a bias towards CBT prior to coding, which could have impacted their ratings (Raue, Putterman, Goldfried, & Wolitzky, 1995).

The coder questionnaires pointed to some potential difficulties with coding the

WAI-O which may have confounded alliance. The coders were inexperienced with psychotherapy generally, which may have made it more difficult for them to code CfD, which has a less mechanised approach than CBT (Hayes, 2004). Though several factors were identified that may have influenced alliance ratings, this is the first piece of research to investigate issues in observer coding of the alliance and points to potential areas of interest for future research to address.

There were some limitations of the sampling methodology employed. The sample size was small and underpowered for hypothesis one by at least 24 participants (though some significant findings were still found). For hypothesis two, the division of the sample by therapy for the regression meant that the sample size was much smaller than originally intended. Resource limitations meant that not many more than 144 hours of coding could have been completed. However, perhaps reducing the number of sessions coded per participant from three to two might have been a better compromise between adequate power and adequate generalisability coefficients. This would have allowed adequate power of $n = 64$ to have been achieved with very little additional coding, while potentially dropping a session (session one) that coders reported was difficult to code for CBT.

Additional factors that may have impacted the alliance-outcome relationship, such as therapist characteristics and dropout could not be examined due to the small sample size. These factors significantly limit the conclusions that can be drawn from the findings. In addition, the sample was not truly randomised, as participants were selected so as to ensure that there was a spread of therapists were included in the sample. However, the findings from the demographic data suggest that this did not lead to any patterns of difference between the two therapies.

Recommendations for Future Research

Further research comparing alliance measures on the same sample for differing

therapies would help to answer some of the questions about the suitability of the WAI for experiential therapies. In addition, in depth qualitative analysis of the experience of coding from the perspective of observers and of therapists and clients completing alliance measures could assist with further understanding the alliance concept and its measurement. In addition, given the ceiling effects present in the WAI-O data, validation of the alternative Likert scales, anchor points and instructions developed by (Darchuk et al., 2000) would be of benefit.

Further investigation into the alliance-outcome relationship for CfD with larger sample sizes would help to understand whether alliance is important for outcome in CfD. This would also allow for the investigation of other factors that may influence the relationship, such as therapist level and client level variables using more complex statistical methods, such as multi-level modelling.

Several authors have pointed to the need for a paradigm shift in the measurement and conceptualisation of the alliance (Doran, 2014; Safran & Muran, 2006). The development and validation of measures that address more dynamic alliance processes, such as rupture and repair would assist in the continued development of alliance research.

Conclusions

This study was the first to examine alliance in CfD and to examine the process of coding alliance from the perspective of observers. The comparison of alliance scores showed higher alliance ratings for CBT compared to CfD for the overall score and goal and task subscales medium effect sizes (Cohen, 1988), though not all achieved statistical significance. Though both therapies were effective only alliance scores for CBT showed any relationship to outcome, with a medium effect size. Due to the small sample size it is not possible to conclude that alliance is unimportant for outcome in

CfD, and this is an area that requires further research.

The coder questionnaires highlight several potential confounding variables that may have influenced alliance ratings, including therapy allegiance and difficulties in coding aspects of CfD, such as the Task and Goal subscales. It is not possible to conclude therefore, whether the findings are due to a genuine difference in the alliance between these two therapies or whether the alliance scores were due to either a lack of understanding of CfD on the part of the coders or whether the WAI-O was less appropriate for CfD than CBT. This is an area that would benefit from future research. The coder questionnaires additionally pointed to difficulties with coding that have been raised by critics of current alliance measures, such as their focus on collaboration.

References

Arnow, B. A., Steidtmann, D., Blasey, C., Manber, R., Constantino, M. J., Klein, D. N., . . . Kocsis, J. H. (2013). The relationship between the therapeutic alliance and treatment outcome in two distinct psychotherapies for chronic depression. *Journal of Consulting and Clinical Psychology, 81,* 627-638. doi: 10.1037/a0031530

Austin, P. C., & Steyerberg, E. W. (2015). The number of subjects per variable required in linear regression analyses. *Journal of Clinical Epidemiology, 68,* 627-636. doi: 10.1016/j.jclinepi.2014.12.014

Barber, J. P., Luborsky, L., Crits-Christoph, P., Thase, M. E., Weiss, R., Frank, A., . . . Gallop, R. (1999). Therapeutic alliance as a predictor of outcome in treatment of cocaine dependence. *Psychotherapy Research, 9,* 54-73. doi: 10.1093/ptr/9.1.54

Beck, J. (1995). Cognitive Therapy: Basics and beyond. New York, NY: Guildford Press.

Bordin, E. S. (1979). The generalizability of the psychoanalytic concept of the working alliance. *Psychotherapy: Theory, Research & Practice, 16,* 252-260. doi:10.1037/h0085885

Cicchetti, D. V. (1994). Guidelines, Criteria, and Rules of Thumb for Evaluating Normed and Standardized Assessment Instruments in Psychology. *Psychological Assessment, 6,* 284-290. doi:10.1037/1040-3590.6.4.284

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd Ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cohen, J. (1992). A Power Primer. *Psychological Bulletin, 112,* 155-159. doi:10.1037/0033-2909.112.1.155

Constantino, J., Castonguay, L., & Schut, A. J. (2002). The working alliance: A flagship for the "scientist-practitioner" model in psychotherapy. In G. S. Tyron (Eds.)

Counseling based on process research: Applying what we know (pp. 81-131), Boston: Allyn & Bacon.

Crits-Christoph, P., Gibbons, M. B. C., Hamilton, J., Ring-Kurtz, S., & Gallop, R. (2011). The Dependability of Alliance Assessments: The Alliance-Outcome Correlation Is Larger than You Might Think. *Journal of Consulting and Clinical Psychology, 79*, 267-278. doi:10.1037/a0023668

Darchuk, A., Wang, V., Weibel, D., Fende, J., Anderson, T., Horvath, A. (2001). *Manual for the Working Alliance Inventory Observer Form (WAI-O): Revision IV.* Retrieved 27.10.15 from: http://wai.profhorvath.com/sites/default/files/upload/WAI-O%20Manual%20V4.pdf

Del Re, A. C., Flückiger, C., Horvath, A. O., Symonds, D., & Wampold, B. E. (2012). Therapist effects in the therapeutic alliance-outcome relationship: A restricted-maximum likelihood meta-analysis. *Clinical Psychology Review, 32*, 642-649. doi:10.1016/j.cpr.2012.07.002

DeRubeis, R. J., Brotman, M. A., & Gibbons, C. J. (2005). A conceptual and methodological analysis of the nonspecifics argument. *Clinical Psychology-Science and Practice, 12*, 174-183. doi:10.1093/clipsy/bpi022

Derubeis, R. J., Brotman, M. A., & Gibbons, C. J. (2005). A Conceptual and Methodological Analysis of the Nonspecifics Argument. *Clinical Psychology: Science and Practice, 12*, 174-183. doi:10.1093/clipsy.bpi022

Doran, J. M. (2014). The working alliance: Where have we been, where are we going? *Psychotherapy Research*, 1-18. doi:10.1080/10503307.2014.954153

Doran, J. M., Safran, J. D., Waizmann, K. B., & Muran, J. C. (2012). The alliance negotiation scale: Psychometric construction and preliminary reliability and validity analysis. *Psychotherapy Research, 22*, 710-719. doi:

10.1080/10503307.2012.709326

Elvins, R., & Green, J. (2008). The conceptualization and measurement of therapeutic alliance: An empirical review. *Clinical Psychology Review, 28*, 1167-1187. doi:10.1016/j.cpr.2008.04.002

Evans, C., Mellor-Clark, J., Margison, F., Barkham, M., Audin, K., Connell, J., & McGrath, G. (2000). CORE: Clinical Outcomes in Routine Evaluation. *Journal of Mental Health, 9*(3), 247-255. doi:10.1080/jmh.9.3.247.255

Fluckiger, C., Del Re, A. C., Wampold, B. E., Symonds, D., & Horvath, A. O. (2012). How central is the alliance in psychotherapy? A multilevel longitudinal meta-analysis. *Journal of Counseling Psychology, 59*, 10-17. doi:10.1037/a0025749

Green, S. B. (1991). How many subjects does it take to do a regression analysis. *Multivariate Behavioral Research, 26*, 499-510. doi:10.1207/s15327906mbr2603_7

Hanson, W. E., Curry, K. T., & Bandalos, D. L. (2002). Reliability Generalization of Working Alliance Inventory Scale Scores. *Educational and Psychological Measurement, 62*, 659-673. doi:10.1177/0013164402062004008

Hayes, S. C. (2004). Acceptance and commitment therapy, relational frame theory, and the third wave of behavioral and cognitive therapies. *Behavior Therapy, 35*, 639-665. doi:10.1016/S0005-7894(04)80013-3

Health and Social Care Information Centre (2016). *Improving access to psychological therapies (IAPT) executive summary.* Retrieved 28.10.17 from: http://Improving-Access-to-Psychological-Therapies-Executive-Summary-April-2016%2FAny%2FIAPT-month-Apr-2016-exec-sum&usg=AOvVaw3kxxm1WXGkriqGt-nzj-xH

Horvath, A. O., & Bedi, R. P. (2002). The alliance. In J. C. Norcross (Eds.), *Psychotherapy relationships that work* (pp. 37-69), New York, NY: Oxford

University Press.

Horvath, A. O., Del Re., C., Fluckiger, C., & Symonds, D., (2011). The alliance. In J. C. Norcross (Eds.), *Psychotherapy relationships that work* (pp. 25-69), New York, NY: Oxford University Press.

Horvath, A. O., Del Re, A. C., Flückiger, C., & Symonds, D. (2011). Alliance in Individual Psychotherapy. *Psychotherapy, 48*, 9-16. doi:10.1037/a0022186

Horvath, A. O., & Greenberg, L. S. (1989). Development and Validation of the Working Alliance Inventory. *Journal of Counseling Psychology, 36,* 223-233. doi:10.1037/0022-0167.36.2.223

Horvath, A. O., & Luborsky, L. (1993). The Role of the Therapeutic Alliance in Psychotherapy. *Journal of Consulting and Clinical Psychology, 61*, 561-573. doi:10.1037/0022-006X.61.4.561

Horvath, A. O., & Marx, R. W. (1990). The development and decay of the working alliance during time-limited counselling. *Canadian Journal of Counselling, 24,* 240-260. Retrieved 28.09.17 from: http://psycnet.apa.org/record/1991-13245-001

Horvath, A. O., & Symonds, B. D. (1991). Relation Between Working Alliance and Outcome in Psychotherapy: A Meta-Analysis. *Journal of Counseling Psychology, 38,* 139-149. doi:10.1037/0022-0167.38.2.139

Kadera, S. W., Lambert, M. J., & Andrews, A. A. (1996). How much therapy is really enough? A session-by-session analysis of the psychotherapy dose-effect relationship. *Journal of Psychotherapy Practice and Research, 5*, 132-151. Retrieved 20.07.17 from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3330412/pdf/132.pdf

Kazdin, A. E., Marciano, P. L., & Whitley, M. K. (2005). The therapeutic alliance in cognitive-behavioral treatment of children referred for oppositional, aggressive,

and antisocial behavior. *Journal of Consulting and Clinical Psychology, 73*, 726-730. doi:10.1037/0022-006X.73.4.726

Kivlighan, D. M., & Shaughnessy, P. (2000). Patterns of working alliance development: A typology of client's working alliance ratings. *Journal of Counseling Psychology, 47*, 362-371. doi:10.1037/0022-0167.47.3.362

Klein, D. N., Schwartz, J. E., Santiago, N. J., Vivian, D., Vocisano, C., Castonguay, L. G., . . . Keller, M. B. (2003). Therapeutic alliance in depression treatment: controlling for prior change and patient characteristics. *Journal of Consulting and Clinical Psychology, 71*, 997-1006. doi: 10.1037/0022-006x.71.6.997

Klug, G., Zimmermann, J., & Huber, D. (2016). Outcome trajectories and mediation in psychotherapeutic treatments of major depression. *Journal of the American Psychoanalytic Association, 64*, 307-343. doi:10.1177/0003065116644742

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ- 9. *Journal of General Internal Medicine, 16*, 606-613. doi:10.1046/j.1525-1497.2001.016009606.x

Krupnick, J. L., Sotsky, S. M., Simmens, S., Moyer, J., Elkin, I., Watkins, J., & Pilkonis, P. A. (1996). The role of the therapeutic alliance in psychotherapy and pharmacotherapy outcome: findings in the National Institute of Mental Health Treatment of Depression Collaborative Research Program. *Journal of consulting and clinical psychology, 64*, 532-539. doi:10.1037//0022-006x.64.3.532

Layard, R., Clark, D., Bell, S., Knapp, M., Meacher, B., Priebe, S., Turnberg, L., Thornicroft, G., & Wright, B. (2006). The depression report: A new deal for depression and anxiety disorders. *The Centre for Economic Performance's Mental Health Policy Group*, LSE.

Leibert, T. W., Smith, J. B., & Agaskar, V. R. (2011). Relationship between the working alliance and social support on counseling outcome. *Journal of clinical*

psychology, 67, 709. doi:10.1002/jclp.20800

Lewis, G. (1994). Assessing psychiatric disorder with a human interviewer or a computer. *Journal of Epidemiology and Community Health, 48,* 207-210. doi: 10.1136/jech.48.2.207

Levene, H. (1960). Robust tests for equality of variances. In I. Olkin, H. Hotelling (Eds.), *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling* (pp. 278-292), Stanford, CA: Stanford University Press.

Lorenzo-Luaces, L., DeRubeis, R. J., & Webb, C. A. (2014). Client characteristics as moderators of the relation between the therapeutic alliance and outcome in cognitive therapy for depression. *Journal of Consulting and Clinical Psychology, 82,* 368-373. doi: 10.1037/a0035994

Marmar, C. R., Gaston, L., Gallagher, D., & Thompson, L. W. (1989). Alliance and outcome in late-life depression. *Journal of Nervous and Mental Disease, 177,* 464-472. doi: 10.1097/00005053-198908000-00003

Marmar, C. R., Weiss, D. S., & Gaston, L. (1989). Toward the validation of the California Therapeutic Alliance Rating System. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 1,* 46-52. doi:10.1037/1040-3590.1.1.46

Martin, A., Rief, W., Klaiberg, A., & Braehler, E. (2006). Validity of the Brief Patient Health Questionnaire mood scale (PHQ-9) in the general population. *General Hospital Psychiatry, 28,* 71-77. doi:10.1016/j.genhosppsych.2005.07.003

Martin, D. J., Garske, J. P., & Davis, M. K. (2000). Relation of the therapeutic alliance with outcome and other variables: a meta-analytic review. *Journal of Consulting and Clinical Psychology, 68,* 438-450. doi:10.1037/0022-006X.68.3.438

McGraw, K. O., & Wong, S. P. (1996). Forming Inferences About Some Intraclass Correlation Coefficients. *Psychological Methods, 1,* 30-46. doi:10.1037/1082-

989X.1.1.30

National Institute for Clinical Excellence (2009). *Depression: The treatment and management of depression in adults (CG90).* Retrieved 17.03.14 from: http://www.nice.org.uk/nicemedia/pdf/CG90NICEguideline.pdf

National Institute for Clinical Excellence (2011). *Generalised anxiety disorder and panic disorder (with or without agoraphobia) in adults: Management in primary, secondary and community care (CG113).* Retrieved 17.03.14 from: http://www.nice.org.uk/nice media/live/ 13314/52601/52601.pdf

Norcross, J. (2002). *Psychotherapy relationships that work: therapist contributions and responsiveness to patients*. New York: Oxford University Press.

Norcross, J. C., & Wampold, B. E. (2011). Evidence-based therapy relationships: research conclusions and clinical practices. *Psychotherapy, 48*, 98-102. doi:10.1037/a0022161

Orlinsky, D. E., & Howard, K. I. (1967). The good therapy hour: Experiential correlates of patients' and therapists' evaluations of therapy sessions. *Archives of General Psychiatry, 16*, 621-632. doi:10.1001/archpsyc.1967.01730230105013

Preschl, B., Maercker, A., & Wagner, B. (2011). The working alliance in a randomized controlled trial comparing online with face-to-face cognitive-behavioral therapy for depression. *Bmc Psychiatry, 11*. doi:10.1186/1471-244x-11-189

Raue, P. J., Goldfried, M. R., & Barkham, M. (1997). The therapeutic alliance in Psychodynamic–Interpersonal and Cognitive–Behavioral Therapy. *Journal of Consulting and Clinical Psychology, 65*, 582-587. doi:10.1037/0022-006X.65.4.582

Raue, P. J., Goldfried, M. R., & Barkham, M. (1997). The therapeutic alliance in psychodynamic-interpersonal and cognitive-behavioral therapy. *Journal of Consulting and Clinical Psychology, 65*, 582-587. doi: 10.1037/0022-

006X.65.4.582

Razali, N., M., & Wah., Y. B. (2011). Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. *Journal of Statistical Modeling and Analytics*, 2, 21-33. Retrieved 9.08.17 from: https://www.researchgate.net/profile/Bee_Yap/publication/267205556_Power_Comparisons_of_Shapiro-Wilk_Kolmogorov-Smirnov_Lilliefors_and_Anderson-Darling_Tests/links/5477245b0cf29afed61446e1/Power-Comparisons-of-Shapiro-Wilk-Kolmogorov-Smirnov-Lilliefors-and-Anderson-Darling-Tests.pdf doi:10.1177/0362153713499545

Rogers (1965). *Client centred therapy*. Cambridge: MA, Houghton Mifflin.

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin, 86,* 638-641. doi:10.1037/0033-2909.86.3.638

Sanders, P., & Hill, A., (2014). *Counselling for depression: A person centred and experiential approach to practice.* London: Sage.

Safran, J. D., & Muran, J. C. (2006). Has the concept of the therapeutic alliance outlived its usefulness? *Psychotherapy: Theory, Research, Practice, Training, 43,* 286-291. doi:10.1037/0033-3204.43.3.286

Saxon, D., Ashley, K., Bishop - Edwards, L., Connell, J., Harrison, P., Ohlsen, S., . . . Barkham, M. (2017). A pragmatic randomised controlled trial assessing the non-inferiority of counselling for depression versus cognitive-behaviour therapy for patients in primary care meeting a diagnosis of moderate or severe depression (PRaCTICED): Study protocol for a randomised controlled trial. *Trials*. doi: 10.1186/s13063-017-1834-6

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika,* 52, 591-611. doi: 10.1093/biomet/52.3-4.591

Shirk, S. R., & Karver, M. (2003). Prediction of treatment outcome from relationship

variables in child and adolescent therapy: A meta-analytic review. *Journal of Consulting and Clinical Psychology, 71*, 452-464. doi:10.1037/0022-006X.71.3.452

Siev, J., Huppert, J., & Chambless, D.L. (2009). The dodo bird, treatment technique, and disseminating empirically supported treatments. *Association of advancement of behaviour therapy, 32*, 69-76.

Snippe, E., Fleer, J., Tovote, K. A., Sanderman, R., Emmelkamp, P. M. G., & Schroevers, M. J. (2015). The Therapeutic Alliance Predicts Outcomes of Cognitive Behavior Therapy but Not of Mindfulness-Based Cognitive Therapy for Depressive Symptoms. *Psychotherapy and Psychosomatics, 84*, 314-315. doi:10.1159/000379755

Strunk, D. R., Cooper, A. A., Ryan, E. T., DeRubeis, R. J., & Hollon, S. D. (2012). The Process of Change in Cognitive Therapy for Depression when Combined with Antidepressant Medication: Predictors of Early Intersession Symptom Gains. *Journal of Consulting and Clinical Psychology, 80*, 730-738. doi:10.1037/a0029281

Tichenor, V. U., & Hill, C. E. (1989). A comparison of six measures of the working alliance. *Psychotherapy, 26*, 195-199. doi: 10.1037/h0085419

Urbaniak, G. C., & Plous, S. (2013). Research Randomizer (Version 4.0) [Computer software]. Retrieved: 1.3.17 from http://www.randomizer.org/

Watson, J. C., & Greenberg, L. S. (2000). Alliance ruptures and repairs in experiential therapy. *Journal of clinical psychology, 56*, 175. doi: 10.1002/(sici)1097-4679(200002)56:2<175::aid-jclp4>3.0.co;2-5

Webb, C. A., DeRubeis, R. J., Amsterdam, J. D., Shelton, R. C., Hollon, S. D., & Dimidjian, S. (2011). Two aspects of the therapeutic alliance: differential relations with depressive symptom change. *Journal of Consulting and Clinical*

Psychology, 79, 279-283. doi:10.1037/a0023252

Webb, C. A., DeRubeis, R. J., & Barber, J. P. (2010). Therapist adherence/competence and treatment outcome: A meta-analytic review. *Journal of Consulting and Clinical Psychology, 78,* 200-211. doi:10.1037/a0018912

Weck, F., Grikscheit, F., Hofling, V., Kordt, A., Hamm, A. O., Gerlach, A. L., . . . Lang, T. (2016). The role of treatment delivery factors in exposure-based cognitive behavioral therapy for panic disorder with agoraphobia. *Journal of Anxiety Disorders, 42,* 10-18. doi:10.1016/j.janxdis.2016.05.007

Zuroff, D. C., Koestner, R., Moskowitz, D. S., McBride, C., Marshall, M., & Bagby, R. M. (2007). Autonomous motivation for therapy: A new common factor in brief treatments for depression. *Psychotherapy Research, 17*, 137-148. doi:10.1080/10503300600919380

Appendices

Appendix E: Ethical Approval

NHS
Health Research Authority
NRES Committee Yorkshire & The Humber - South Yorkshire
North East REC Centre
Unit 002, TEDCO Business Centre
Rolling Mill Road
Jarrow
Tyne and Wear
NE32 3DT

Telephone: 0191 428 3561

27 March 2014

Professor Michael Barkham
Director, Centre for Psychological Services Research
University of Sheffield
Dept of Psychology
University of Sheffield
Western Bank
SHEFFIELD
S10 2TN

Dear Professor Barkham

Study title:	**A pragmatic non-inferiority randomised controlled trial of the clinical and cost-effectiveness of counselling for depression versus cognitive-behaviour therapy, for clients in primary care meeting a diagnosis of moderate or severe depression: The PRaCTICED Trial**
REC reference:	**14/YH/0001**
IRAS project ID:	**130352**

Thank you for your letter of 25 March 2014. I can confirm the REC has received the documents listed below and that these comply with the approval conditions detailed in our letter dated 04 February 2014

Documents received

The documents received were as follows:

Document	*Version*	*Date*
Interview Schedules/Topic Guides	Interview Topic Guide / Version 2.0	25 March 2014
Participant Consent Form: Main Consent Form	Version 2.0	25 March 2014

Appendix E: Ethical Approval Continued

Approved documents

The final list of approved documentation for the study is therefore as follows:

Document	*Version*	*Date*
Covering Letter	Michael Barkham	11 December 2013
Evidence of insurance or indemnity	The University of Sheffield	13 November 2013
GP/Consultant Information Sheets	Information Sheet for GP, V1.0	20 November 2013
GP/Consultant Information Sheets	GP Notification, V1.0	20 November 2013
GP/Consultant Information Sheets	GP Risk Letter, V1.0	20 November 2013
Interview Schedules/Topic Guides	Brief Exit Interview, V1.0	20 November 2013
Interview Schedules/Topic Guides	Interview Topic Guide / Version 2.0	25 March 2014
Investigator CV	Michael Barkham	16 November 2013
Letter of invitation to participant	Patient Information from PWP, V1.0	20 November 2013
Other: CV - Student Research Supervisor	Gillian E. Hardy	16 November 2013
Other: Student CV	Caroline Dunsmuir-White	16 November 2013
Other: Student CV	Kim Campbell	16 November 2013
Other: Letter from Funder	BACP Research Foundation	19 November 2013
Other: Adverse Events Reporting - Study Specific Procedure	V1.0	20 November 2013
Other: Risk Protocol - Research Interviews	V1.0	20 November 2013
Other: Risk Form for Suicide and Self Harm	V1.0	20 November 2013
Other: Risk Form for Risks Not Including Suicide and Self Harm	V1.0	20 November 2013
Participant Consent Form: Consent to Contact	V1.0	20 November 2013
Participant Consent Form: Consent to Interview	V1.0	20 November 2013
Participant Consent Form: Main Consent Form	Version 2.0	25 March 2014
Participant Information Sheet: PIS Main Trial	V1.0	20 November 2013
Participant Information Sheet: Patient Information for Assessment	V1.0	20 November 2013
Protocol	V1.0	20 November 2013
Questionnaire: Treatment Preference		
Questionnaire: Therapy Expectation Form		
Questionnaire: CIS-R		
Questionnaire: BDI-II		
Questionnaire: PHQ-9		
Questionnaire: GAD-7		
Questionnaire: EQ-5D		
Questionnaire: Wellbeing-VAS		
Questionnaire: CORE-OM		
Questionnaire: Work and Social Adjustment Scale		
Questionnaire: CD-RISC		

Appendix E: Ethical Approval Continued

Questionnaire: MINI Diagnostic - Sections I&J		
Questionnaire: CSSRI-EU		
Questionnaire: Client Satifsfaction		
REC application	IRAS V3.5	21 November 2013
Referees or other scientific critique report	Peer Review 1	
Referees or other scientific critique report	Peer Review 2	
Referees or other scientific critique report	Peer Review 3	
Summary/Synopsis	Flowchart Recruitment, V1.0	20 November 2013
Summary/Synopsis	Consort, Wave 1, V1.0	20 November 2013
Summary/Synopsis	Consort: Wave 2, V1.0	20 November 2013
Summary/Synopsis	SOP Recruitment	20 November 2013
Summary/Synopsis	SOP Patient Treatment, V1.0	20 November 2013

You should ensure that the sponsor has a copy of the final documentation for the study. It is the sponsor's responsibility to ensure that the documentation is made available to R&D offices at all participating sites.

14/YH/0001 **Please quote this number on all correspondence**

Yours sincerely

Kerry Dunbar
REC Assistant

E-mail: nrescommittee.yorkandhumber-southyorks@nhs.net

Copy to: *Mr David Saxon, University of Sheffield*
Mr Nicolas Bell, Sheffield Health & Social Care NHS Foundation Trust

Appendix E: Ethical Approval Continued

NHS
Health Research Authority
NRES Committee Yorkshire & the Humber - South Yorkshire
North East REC Centre
Unit 002, TEDCO Business Centre
Rolling Mill Road
Jarrow
Tyne and Wear
NE32 3DT

Telephone: 0191 428 3566
Facsimile: 0191 428 3432

04 February 2014

Professor Michael Barkham
Director, Centre for Psychological Services Research
Department of Psychology
University of Sheffield
Western Bank
SHEFFIELD
S10 2TN

Dear Professor Barkham

Study title:	**A pragmatic non-inferiority randomised controlled trial of the clinical and cost-effectiveness of counselling for depression versus cognitive-behaviour therapy, for clients in primary care meeting a diagnosis of moderate or severe depression: The PRaCTICED Trial**
REC reference:	**14/YH/0001**
IRAS project ID:	**130352**

The Research Ethics Committee reviewed the above application at the meeting held on the 30 January 2014. Thank you for attending to discuss the application.

We plan to publish your research summary wording for the above study on the HRA website, together with your contact details, unless you expressly withhold permission to do so. Publication will be no earlier than three months from the date of this favourable opinion letter. Should you wish to provide a substitute contact point, require further information, or wish to withhold permission to publish, please contact the REC Manager Mrs Joan Brown, nrescommittee.yorkandhumber-southyorks@nhs.net.

Ethical opinion

It was queried whether you were applying for approval of the whole RCT as well as what the students would be doing and you confirmed that ethical approval was being sought for the whole trial.

It was observed that the only issue with the application was that there was no indication of the topics that would be discussed with the people who dropped out of the study. It was explained that this was a work in progress and would be submitted to the REC once it had been finalised.

Appendix E: Ethical Approval Continued

It was observed there was a minor clarification required in the consent form.

The members of the Committee present gave a favourable ethical opinion of the above research on the basis described in the application form, protocol and supporting documentation, **subject to the conditions specified below.**

Ethical review of research sites

NHS Sites

The favourable opinion applies to all NHS sites taking part in the study, subject to management permission being obtained from the NHS/HSC R&D office prior to the start of the study (see "Conditions of the favourable opinion" below).

Conditions of the favourable opinion

The favourable opinion is subject to the following conditions being met prior to the start of the study.

1. **Submit a revised Consent Form as follows: Amend Point 5 to read "I understand that data collected during the study may be looked at by individuals from the study team or individuals from regulatory authorities or the NHS Trust where it is relevant to my taking part in this study. I give permission for these individuals to have access to my records"**

2. **Submit a copy of the interview schedule that will be used for people who drop out of the study once it has been finalised for information only. There is no need for the schedule to be approved by the REC.**

You should notify the REC in writing once all conditions have been met (except for site approvals from host organisations) and provide copies of any revised documentation with updated version numbers. The REC will acknowledge receipt and provide a final list of the approved documentation for the study, which can be made available to host organisations to facilitate their permission for the study. Failure to provide the final versions to the REC may cause delay in obtaining permissions.

Management permission or approval must be obtained from each host organisation prior to the start of the study at the site concerned.

Management permission ("R&D approval") should be sought from all NHS organisations involved in the study in accordance with NHS research governance arrangements.

Guidance on applying for NHS permission for research is available in the Integrated Research Application System or at http://www.rdforum.nhs.uk.

Where a NHS organisation's role in the study is limited to identifying and referring potential participants to research sites ("participant identification centre"), guidance should be sought from the R&D office on the information it requires to give permission for this activity.

For non-NHS sites, site management permission should be obtained in accordance with the procedures of the relevant host organisation.

Sponsors are not required to notify the Committee of approvals from host organisations

Appendix E: Ethical Approval Continued

Registration of Clinical Trials

All clinical trials (defined as the first four categories on the IRAS filter page) must be registered on a publically accessible database within 6 weeks of recruitment of the first participant (for medical device studies, within the timeline determined by the current registration and publication trees).

There is no requirement to separately notify the REC but you should do so at the earliest opportunity e.g when submitting an amendment. We will audit the registration details as part of the annual progress reporting process.

To ensure transparency in research, we strongly recommend that all research is registered but for non clinical trials this is not currently mandatory.

If a sponsor wishes to contest the need for registration they should contact Catherine Blewett (catherineblewett@nhs.net), the HRA does not, however, expect exceptions to be made. Guidance on where to register is provided within IRAS.

It is responsibility of the sponsor to ensure that all the conditions are complied with before the start of the study or its initiation at a particular site (as applicable).

Approved documents

The documents reviewed and approved at the meeting were:

Document	*Version*	*Date*
Covering Letter	Michael Barkham	11 December 2013
Evidence of insurance or indemnity	The University of Sheffield	13 November 2013
GP/Consultant Information Sheets	Information Sheet for GP, V1.0	20 November 2013
GP/Consultant Information Sheets	GP Notification, V1.0	20 November 2013
GP/Consultant Information Sheets	GP Risk Letter, V1.0	20 November 2013
Interview Schedules/Topic Guides	Interview Topic Guide, V1.0	20 November 2013
Interview Schedules/Topic Guides	Brief Exit Interview, V1.0	20 November 2013
Investigator CV	Michael Barkham	16 November 2013
Letter of invitation to participant	Patient Information from PWP, V1.0	20 November 2013
Other: CV - Student Research Supervisor	Gillian E. Hardy	16 November 2013
Other: Student CV	Caroline Dunsmuir-White	16 November 2013
Other: Student CV	Kim Campbell	16 November 2013
Other: Letter from Funder	BACP Research Foundation	19 November 2013
Other: Adverse Events Reporting - Study Specific Procedure	V1.0	20 November 2013
Other: Risk Protocol - Research Interviews	V1.0	20 November 2013
Other: Risk Form for Suicide and Self Harm	V1.0	20 November 2013
Other: Risk Form for Risks Not Including Suicide and Self Harm	V1.0	20 November 2013

Appendix E: Ethical Approval Continued

Participant Consent Form: Main Consent Form	V1.0	20 November 2013
Participant Consent Form: Consent to Contact	V1.0	20 November 2013
Participant Consent Form: Consent to Interview	V1.0	20 November 2013
Participant Information Sheet: PIS Main Trial	V1.0	20 November 2013
Participant Information Sheet: Patient Information for Assessment	V1.0	20 November 2013
Protocol	V1.0	20 November 2013
Questionnaire: Treatment Preference		
Questionnaire: Therapy Expectation Form		
Questionnaire: CIS-R		
Questionnaire: BDI-II		
Questionnaire: PHQ-9		
Questionnaire: GAD-7		
Questionnaire: EQ-5D		
Questionnaire: Wellbeing-VAS		
Questionnaire: CORE-OM		
Questionnaire: Work and Social Adjustment Scale		
Questionnaire: CD-RISC		
Questionnaire: MINI Diagnostic - Sections I&J		
Questionnaire: CSSRI-EU		
Questionnaire: Client Satisfaction		
REC application	IRAS V3.5	21 November 2013
Referees or other scientific critique report	Peer Review 1	
Referees or other scientific critique report	Peer Review 2	
Referees or other scientific critique report	Peer Review 3	
Summary/Synopsis	Flowchart Recruitment, V1.0	20 November 2013
Summary/Synopsis	Consort, Wave 1, V1.0	20 November 2013
Summary/Synopsis	Consort: Wave 2, V1.0	20 November 2013
Summary/Synopsis	SOP Recruitment	20 November 2013
Summary/Synopsis	SOP Patient Treatment, V1.0	20 November 2013

Membership of the Committee

The members of the Ethics Committee who were present at the meeting are listed on the attached sheet.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

Appendix E: Ethical Approval Continued

After ethical review

Reporting requirements

The attached document "After ethical review – guidance for researchers" gives detailed guidance on reporting requirements for studies with a favourable opinion, including:

- Notifying substantial amendments
- Adding new sites and investigators
- Notification of serious breaches of the protocol
- Progress and safety reports
- Notifying the end of the study

The NRES website also provides guidance on these topics, which is updated in the light of changes in reporting requirements or procedures.

Feedback

You are invited to give your view of the service that you have received from the National Research Ethics Service and the application procedure. If you wish to make your views known please use the feedback form available on the website.

Further information is available at National Research Ethics Service website > After Review

14/YH/0001	**Please quote this number on all correspondence**

We are pleased to welcome researchers and R & D staff at our NRES committee members' training days – see details at http://www.hra.nhs.uk/hra-training/

With the Committee's best wishes for the success of this project.

Yours sincerely

J Brown

***pp* Ms Jo Abbott**
Chair

Email: nrescommittee.london-camdenandislington@nhs.net

Enclosures: *List of names and professions of members who were present at the meeting and those who submitted written comments*
"After ethical review – guidance for researchers" SL-AR-2

Copy to: *Mr David Saxon, University of Sheffield*

Mr Nicolas Bell, Sheffield Health & Social Care NHS Foundation Trust

Appendix E: Ethical Approval Continued

NRES Committee Yorkshire & the Humber - South Yorkshire

Attendance at Committee meeting on 30 January 2014

Committee Members:

Name	*Profession*	*Present*	*Notes*
Ms Jo Abbott (Chair)	Consultant in Public Health	Yes	
Dr Ahmed H Abdelhafiz	Consultant Physician, Elderly Medicine	Yes	
Dr Peter Allmark	Principal Nursing Lecturer	No	
Reverend Joan Ashton	Co-ordinator of Chaplaincy Services	Yes	
Ms Helen Barlow	Knowledge Service Manager	Yes	
Professor Nigel Beail	Consultant Clinical Psychologist & Professor of Psychology	Yes	
Mr Ian Cawthorne	Chief Pharmacist	No	
Ms Susan Hampshaw	Head of Research, Evaluation and Innovation	Yes	
Mr Neil Marsden	Police Staff	Yes	
Dr Duane Mellor	Lecturer in Dietetics	No	
Mrs Andrea Porritt	Community Specialist Practitioner/District Nurse	Yes	
Mrs Carole Taylor	Deputy Chief Pharmacist	Yes	

Also in attendance:

Name	*Position (or reason for attending)*
Ms Joan Brown	REC Manager

Appendix E: Ethical Approval Continued

The University Of Sheffield.

Department Of Psychology.
Clinical Psychology Unit.

Doctor of Clinical Psychology (DClin Psy) Programme
Clinical supervision training and NHS research training & consultancy.

Clinical Psychology Unit
Department of Psychology
University of Sheffield
Western Bank
Sheffield S10 2TP UK

Telephone: 0114 2226570
Fax: 0114 2226610
Email: dclinpsy@sheffield.ac.uk
Please address any correspondence to Ian Macdonald, Research Support Officer

22nd May 2015

To: Research Governance Office

Dear Sir/Madam,

RE: Confirmation of Scientific Approval and indemnity of enclosed Research Project

Project title: The Relationship Between Working Alliance and Premature Termination from Psychotherapy: Results from the PRaCTICED trial
Investigators: Elizabeth Gilley (DClin Psy Trainee, University of Sheffield); Professor Gillian Hardy (Academic Supervisor, University of Sheffield).

I write to confirm that the enclosed proposal forms part of the educational requirements for the Doctoral Clinical Psychology Qualification (DClin Psy) run by the Clinical Psychology Unit, University of Sheffield.

Three independent reviewers appointed by the Clinical Psychology Unit Research Sub-committee have scientifically reviewed it.

I can confirm that all necessary amendments have been made to the satisfaction of the reviewers, who are now happy that the proposed study is of sound scientific quality. Consequently, the University will also indemnify it and would be happy to act as research sponsor once ethical approval has been gained.

Given the above, I would remind you that the Unit already has an agreement with your office to exempt this proposal from further scientific review. However, if you require any further information, please do not hesitate to contact me.

Yours sincerely

Dr. Andrew Thompson
Director of Research Training

Cc. Elizabeth Gilley; Professor Gillian Hardy

Appendix E: Ethical Approval Continued

The University Of Sheffield.

Department Of Psychology.
Clinical Psychology Unit.

Doctor of Clinical Psychology (DClin Psy) Programme
Clinical supervision training and NHS research training & consultancy.

Clinical Psychology Unit
Department of Psychology
University of Sheffield
Western Bank
Sheffield S10 2TN UK

Telephone: 0114 22 26650
Fax: 0114 22 26610
Email: ian.macdonald@sheffield.ac.uk

8th June 2015
Elizabeth Gilley
Trainee Clinical Psychologist
Department of Psychology
Western Bank

Project title: The Relationship Between Working Alliance and Premature Termination from Psychotherapy: Results from the PRaCTICED trial

6 digit URMS number: 144768

Dear Elizabeth Gilley,

LETTER TO CONFIRM THAT THE UNIVERSITY OF SHEFFIELD IS THE PROJECT'S RESEARCH GOVERNANCE SPONSOR

The University has reviewed the following documents:

1. A University approved URMS costing record;
2. Confirmation of independent scientific approval;
3. Confirmation of independent ethics approval.

All the above documents are in place. Therefore, the University now **confirms** that it is the project's research governance sponsor and, as research governance sponsor, **authorises** the project to commence any non-NHS research activities. Please note that NHS R&D approval will be required before the commencement of any activities which do involve the NHS.

You are expected to deliver the research project in accordance with the University's policies and procedures, which includes the University's Good Research & Innovation Practices Policy: www.shef.ac.uk/ris/other/gov-ethics/grippolicy, Ethics Policy: www.sheffield.ac.uk/ris/other/gov-ethics/ethicspolicy and Data Protection Policies: www.shef.ac.uk/cics/records

Your Supervisor, with your support and input, is responsible for monitoring the project on an ongoing basis. Your Head of Department is responsible for independently monitoring the project as appropriate. The project may be audited during or after its lifetime by the University. Monitoring responsibilities are listed in Annex 1.

Yours sincerely

Dr Andrew Thompson
Director of Research Training, Clinical Psychology Unit

Cc: Professor Gillian Hardy (supervisor);
Professor Paul Overton (Head of Department).

Appendix F: Participant Information Sheet

Information about the research

PRaCTICED Study

A randomised trial comparing the effectiveness of cognitive behavior therapy and counselling for depression

Thank you very much for agreeing to be contacted about the above research study. This information sheet explains the purpose of the study and what will happen if you take part. Please contact us if anything is not clear and talk to others about the study if you wish. You will have a further opportunity to discuss the study with researchers before consenting to full involvement.

What is the purpose of the study?

Depression is a common problem that affects many people and can sometimes be hard to manage. Experts recommend that people with depression receive a 'talking treatment' and/or medication. Your GP may have prescribed some medication for you but this is not always enough on its own. This is where talking therapies can be very helpful.

There are different forms of talking treatments. Our research is trying to find out whether there is a difference between two particular approaches in the treatment of depression: Cognitive Behaviour Therapy (CBT) or Counselling for Depression (CfD).

- Counselling for Depression (CfD) aims to address depression by providing the opportunity for clients to talk about underlying feelings. The therapist and client work together to make personal sense of these feelings.
- Cognitive Behaviour Therapy (CBT) looks at how we think about a situation and how this affects the way we act. The therapist and client work together in changing the client's behaviours, or their thinking patterns, or both of these.

The Sheffield IAPT service delivers both these treatments in its routine service. The purpose of this trial will be to see if there are differences between these two treatments and whether some people are more suited to one form of treatment rather than the other. The study will also tell us what it is about the treatments that people like or dislike so that we can improve them for other people.

Both treatments will be for a minimum of 8 sessions and will normally be for up to 16 sessions but can be up to 20 sessions. Taking part in the study does not mean that you cannot receive treatment later from the Sheffield service.

Do I have to take part?

It is your decision to take part. If you do agree, we will then ask you to sign a consent form. You are free to withdraw at any time, without giving a reason. Leaving the study will not affect the standard of care you receive. However, it is always helpful to understand why someone leaves treatment, in order to try and improve services. We will not try to change your decision.

What will happen to me if I am willing to take part?

About 3 weeks prior to your therapy starting, a researcher will contact you by your chosen method, to invite you to a one-off assessment interview. This will be based at a location as convenient to you as possible. The invitation to this meeting will include a one-day bus pass in case there is a need to use a bus to attend the meeting. We have done this so that no one is out of pocket for attending this one-off meeting.

At the meeting, you can ask any questions you might have about the study. The researcher will ask you a number of questions that will help to see whether the trial is appropriate. If it is, then you will be informed which treatment you will receive. You stand an equal chance of receiving either treatment. You will then be asked to complete some forms.

You do not have to take part unless you feel completely happy with the study.

What are the treatments?

The treatments are Counselling for Depression (CfD) and Cognitive Behaviour Therapy (CBT) and were briefly described earlier.

Both treatments are psychological therapies that have been recommended by NICE (National Institute Clinical Excellence) for the treatment of depression.

What if I have a very strong preference and don't want to receive one of the treatments?

People may have a preference for one treatment over the other. This is understandable. However, if you have a very strong preference, such that, you would be unwilling to receive one of the treatments if you were given it, then please talk to the assessor. If after talking with them you feel the same, then the assessor will ensure that you are referred back to the normal service without losing your place on the waiting list.

How is it decided who gets which treatment?

Sometimes it is not always clear which is the best way of treating patients To find out, we need to compare different treatments. We allocate people to one of two treatments then compare the results to see if one treatment works better for some people while another works better for others.

To try to make sure patients in each treatment are similar to start with, each patient is allocated a treatment by chance. You will have an equal chance of receiving either cognitive behaviour therapy or counselling for depression.

What else will be involved if I take part?

It is standard practice in this service for the sessions to be audiotaped. This is to enable the person you will be seeing to have regular supervision on their work, this is required by the service to ensure we offer the best service.

For the research, a small number of recordings will be listened to by a researcher in order to check the quality of the talking therapy people are receiving. If they do listen to a tape, it will be under strict confidentiality agreements. Some other tapes will also be used as part of the research in order to increase the understanding about how these talking therapies help people who are experiencing depression.

At six months and 12 months after the meeting with the researcher, we will send you a set of questions to see how you are feeling. These will be similar to those forms completed at the start. The actual research study will take 3 years to complete, but you will only be involved for 12 months.

We will ask patients for permission to contact them by their preferred choice (standard mail, email, phone) if they decide to end treatment. This is for us, as researchers, to understand why this has happened. It is not to try to change your decision. However, if you do not wish to take part at that time, then we will respect that decision.

We will also like to conduct some interviews with some people when they complete their treatment. We will not be interviewing everyone but we need your permission to approach you if you are selected. We will only ask about 1 in 10 patients. You do not have to agree to this and saying 'No' will not affect your involvement in the trial or any treatment in the future.

If you are interested in taking part in the separate interview study, we will provide you with more information before you make the decision.

What are the possible disadvantages and risks of taking part?

Both treatments are used in the routine service, so we are not introducing a new treatment. There are no known side effects of either treatment. We are trying to find out a bit more about what works best for particular people, so we have no reason to believe that any one is being disadvantaged. If you had a strong preference for one treatment, then you will have declared that and the trial would not be appropriate for you.

At any point during the study you can leave without having to give a reason why.

Will I receive any payment for taking part?

We will provide a free one-day bus pass to attend the initial assessment (regardless of whether you have to use it or not). We will also enclose a £10 shopping voucher with the questionnaires at 6-months and 12-months. These will be sent to you regardless of whether you complete the forms or not. However, we hope that this will off set the time spent on completing the forms and very much hope you do.

What happens if new information becomes available during the course of the study?

Sometimes during a study, new information becomes available about the treatment being studied. If this happens, the research team will tell you and discuss whether you want to continue in the study. If you decide to stop taking part in the study your usual care will continue. If you decide to continue in the study you may be asked to sign an updated consent form. If we think you should withdraw from the study, we will explain the reasons and arrange for your care to continue.

What happens when the study stops?

Very occasionally a study is stopped early. If this happens, the reasons will be explained to you and arrangements made for your ongoing care.

What if there is a problem?

If you have a concern about any aspect of this study you should ask to speak to the researcher (Lindsey Bishop-Edwards tel: 07710 388985) or the chief investigator, Michael Barkham (tel: 0114 222 0817) who will do their best to answer your questions.

If they are unable to resolve your concern or you wish to make a complaint regarding the study, please contact the University Research Practice and Governance Co-ordinator Richard Hudson by email to r.j.hudson@sheffield.ac.uk

What will happen to information about me collected during the study?

All information will be held securely and in strict confidence. Only authorised people working on the study will have access to your information and this is kept securely. Where possible, a unique study ID number will be allocated to replace any identifier and only authorised researchers that need to contact you will have access to your personal contact details.

We will destroy all personal details 5 years after the end of the study.

We keep the health information we collect about you separate from your personal

details. We will use the information we collect to look at how best to help people with depression. We will keep it 20 years and then destroy it securely.

Involvement of your GP

We will tell your GP that you are taking part in the study. No other results will be given to your GP.

If we are worried that you are having thoughts about harming yourself, we may need to discuss these with your GP. We will, of course, discuss this with you.

What will happen to the results of the study?

When the study is completed, the results will be published in a scientific journal so that health care professionals can see the results. Your identity and personal details will be kept confidential and no named information about you will be published in any reports.

Who is organising and funding the study?

This study is organised by the University of Sheffield. The funder is the British Association of Counselling and Psychotherapy (BACP) Research Foundation.

Who has reviewed the study?

This study has been reviewed by an independent group of people, called the Research Ethics Committee, to protect your safety, rights, well-being and dignity. The study has been given a favourable opinion by NRES Committee Yorkshire & The Humber - South Yorkshire Ethics committee.

Who is the study co-ordinator?

The study co-ordinator can be contacted by telephone on: (07710 388985). Alternatively, you can write to the researcher at:

PRaCTICED

ScHARR

Regent's Court, 30 Regent's Street

Sheffield, S1 4DA

Email: practiced@sheffield.ac.uk

Thank your for taking time to read this information sheet

Appendix G: Client Consent Form

PRaCTICED Study

Research participant consent form

If you wish to take part in the PRACTICED study, **please place your initials in each of the boxes below, sign and date this form and return it to us in the pre-paid envelope provided.**

Please **INITIAL** box

1	I confirm that I have read and understand the information sheet dated …….. (version ..) for the above study. I have had the opportunity to consider the information, ask questions about the study and understand why this research is being done	☐
2	I agree to an interview with a member of the study team. This will either be face-to-face or by phone and I will be able to choose which one suits me better.	☐
3	I agree to my interview being recorded for the purposes of the research	☐
4	I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason, without my medical care or legal rights being affected	☐

Name of patient (BLOCK CAPITALS) Date Signature

---------- ---------- ----------

Name of person taking consent Date Signature

FOR COMPLETION BY RESEARCHER ONLY **PARTICIPANT ID:** ☐☐☐☐

Appendix H: Trial consent form

PRaCTICED Study

Research participant consent form

If you are interested in taking part in the PRaCTICED study, please read through the points below and note any queries you may have. When you attend the assessment with a member of the research team, they will talk you through the points and answer any questions you may have about the study. Only then will you be asked to complete this form.

Please **INITIAL** box

1	I confirm that I have read and understand the information sheet dated …….. (version ..) for the above study. I have had the opportunity to consider the information, ask questions about the study and understand why this research is being done	
2	I understand that I may not be eligible to take part in the study	
3	I agree to complete the relevant questionnaires at 3, 6 and 12 months after entering the study	
4	I agree to my GP being informed of my participation in the study and of any health concerns the study team may become aware of during my participation	
5	I understand that data collected during the study – as with all data collected within routine NHS service delivery – may be looked at by individuals from the study team or individuals from regulatory authorities or the NHS Trust where it is relevant to my taking part in this study. I give permission for these individuals to have access to my records	
6	I understand that, as part of normal practice, my sessions will be audio-recorded for the purposes of supervision	
7	I understand that some of these audio-recordings may be listened to by researchers either with the purpose of ensuring that the treatments are being delivered appropriately or to enable a better understanding of these treatments	
8	I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason, without my medical care or legal rights being affected	
9	I understand that I may be approached to take part in an additional interview as part of the study, and that I will be given further information and another consent form	
10	I agree to take part in the above study	

Name of patient (BLOCK CAPITALS) Date Signature

---------- ---------- ----------

Name of person taking consent Date Signature

Appendix I:

This page is blank for copyright reasons

Appendix I: WAI-O Form continued

This page is blank for copyright reasons

Appendix I: WAI-O Form continued

This page is blank for copyright reasons

Appendix I: WAI-O Form continued

This page is blank for copyright reasons

Appendix I: WAI-O Form continued

This page is blank for copyright reasons

Appendix J: Copyright Permission WAI

SOCIETY FOR PSYCHOTHERAPY RESEARCH

February 5, 2017

Elizabeth Gilley
University of Sheffield
UK

Dear Ms. Gilley:

You have our permission to use the Working Alliance Inventory (WAI) for your research project. Please be aware that we require publishing the following note at the end of the measure:

Reprinted by permission of the Society for Psychotherapy Research © 2016.

We wish you the best in your work. Please consider joining the Society for Psychotherapy Research, an international, multidisciplinary scientific association devoted to research on psychotherapy. SPR also plays an important role in providing opportunities for interaction and dialogue between researchers and clinicians interested in psychotherapy. You may read more about us at www.psychotherapyresearch.org.

Sincerely,

Marna S. Barrett, Ph.D.
Executive Officer
sprexecutive@gmail.com

Appendix K: PHQ-9

This page is blank for copyright reasons

Appendix L: CORE-OM

This page is blank for copyright reasons

Appendix L: CORE-OM continued

This page is blank for copyright reasons

Appendix M: Coder feedback form

RATER QUESTIONNAIRE
Name: **Date completed:**
Prior to you starting coding, did you have knowledge or experience of any therapies (e.g. CBT, psychodynamic, counselling etc.)? **If yes, were there any therapies that you felt more positive about?**
When using the WAI-O, were there any items that you found particularly difficult to score? If so, please tell me about it.
Were there any times when using the WAI-O that you found that your scoring did not reflect the working alliance that you had observed in the recording? If so, please tell me about it.

Appendix M: Coder Feedback Form continued

Now that you have completed your coding, what can you tell me about the advantages and disadvantages of using the WAI-O ?
Now that you have completed your coding, do you prefer any therapy?
If you were training someone to use the WAI-O, what would be important for them to know?
I consent for the contents of this form to be anonymously used for the purposes of research and future publications. I understand that my words may be directly quoted. Signed: Date:

Appendix N: Training for Coders

CODING USING THE WORKING ALLIANCE INVENTORY – OBSERVER FORM

Liz Gilley
egilley1@sheffield.ac.uk

PLAN FOR THE DAY

1. Introductions
2. The project
3. The working alliance
4. The WAI –O
5. Process for coding

6. Data protection
7. Self care

Appendix N: Training for Coders continued

INTRODUCTIONS

What brings you here?

What are you hoping for from today?

BACKGROUND TO THE PROJECT

"Working Alliance in Cognitive Behaviour Therapy and Counselling for Depression: a Comparison of therapies and Relationship to Outcome"

Does the relationship between client and therapist (WA) differ between therapies?

Does the relationship between client and therapist in psychotherapy impact on the effectiveness of the therapy?

Appendix N: Training for Coders continued

BACKGROUND TO THE PROJECT

HYPOTHESES

1. Working alliance ratings will differ between CBT and CfD
   IV = therapy type DV = WAI-O score
2. Working alliance will be a significant predictor of outcome after controlling for other predictors

Data for the project is being taken from the PRaCTICED trial in Sheffield.

THE ROLE OF THE CODER

You are generating the most important data for the project!

What will it involve?

- Based at SCHaRR
- Listening to hour long therapy recordings (x15)
- Using a psychometric measure (WAI-O) to rate the therapeutic relationship in the recording
- Working with other coders to discuss recordings that are double rated

Appendix N: Training for Coders continued

THE WORKING ALLIANCE

What is the working alliance in psychotherapy?

What do you think makes a good working alliance?

Why is working alliance important?

THE WORKING ALLIANCE

https://www.youtube.com/watch?v=ofgp69DYEks

Appendix N: Training for Coders continued

THE WORKING ALLIANCE

- Working alliance: represents the collaborative parts of the therapeutic relationship - involves both therapist and client (Constantino, Castonguay & Schut, 2002).
- Bordin (1979): definition of working alliance
  - Goal: areas that are targets for change (e.g. reducing symptoms)
  - Task: therapeutic strategies or methods (e.g. exploration, confrontation & direction)
  - Bond: the 'mutual liking, respect and trust between client and therapist' (Raue et al. (n.d.)

WORKING ALLIANCE

What do you notice in these clips?

https://www.youtube.com/watch?v=UpL3ncoK99U

https://www.youtube.com/watch?v=Y7B1IL1ouFw

Appendix N: Training for Coders continued

THE WORKING ALLIANCE INVENTORY

- Measure developed based on Bordin's definition of WA: goal, task and bond by Horvath & Greenberg (1989)
- Originally measures were designed for therapist and client to complete
- WAI-O developed by Tichenor & Hill (1989)
- 36 item measure
- Good internal consistency and interrater reliability

THE WORKING ALLIANCE INVENTORY

Spend the next 15 minutes having a look through the WAI-O (coding pack item 2)and think about the following:

What might be some of the challenges of rating the WAI-O?

Any questions you might have

Appendix N: Training for Coders continued

THE WORKING ALLIANCE INVENTORY: CODING

- Listen to therapy tape in full. Make occasional notes if helpful
- Fill out the cover sheet and measure referring to the instructions for coding:
  - Consider both client and therapist when coding
  - Give the relationship 'benefit of the doubt' – start with most positive score (could be 1 or 7 depending on wording) and subtract or add as required.
  - Do consider your own personal reactions as well as the interaction in therapy e.g. if you feel the therapist isn't being genuine.
  - Refer to more detailed instructions for individual items (coding pack item 4).

THE WORKING ALLIANCE INVENTORY: CODING

CODING PRACTICE UNTIL SCORES ALIGN

Appendix N: Training for Coders continued

PROCESS FOR CODING

1. Get therapy recordings downloaded onto memory stick at SCHaRR
2. Fill out personal details on WAI-O cover sheet.
3. Listen to start of recording and fill out additional details on cover sheet.
4. Listen to recording
5. Complete WAI-O.
6. If doing double rating – fill out double coding form (coding pack item 3).
7. Return completed forms to Liz electronically or to SCHARR.

DATA PROTECTION

Why is data protection important for this project?

Appendix N: Training for Coders continued

DATA PROTECTION: YOUR RESPONSIBILITIES

- Complete online information governance training
- Only accept recordings on an encrypted memory stick
- Use an encrypted computer to listen to recordings
- Only listen to recordings in a private space using headphones
- Do not discuss the contents of what you hear with anyone
- If you recognise/know either client or therapist, stop listening
- Do not put confidential information on forms – initials only
- Keep paper copies locked /password protect online

SELF CARE

Why is self care important for you all?

Appendix N: Training for Coders continued

SELF CARE

Signs of secondary trauma	
Anxiety	Diminished ability
Intrusive thoughts	Irritability
Apathy	Feeling overwhelmed
Depression	Hypervigilance
Lessened enthusiasm	Emotional disturbance
Desensitisation	Disordered thinking

SELF CARE

What can you do to take care of yourself?

Appendix O: Rater Confidentiality Form and Guidelines

Guidance Notes and Confidentiality form for Coders of Therapy Recordings

Doctorate in Clinical Psychology, University of Sheffield

Coding of Therapy Recordings Confidentiality Form & Guidance Notes

Type of project: Research thesis

Project title: Working Alliance in Cognitive Behaviour Therapy and Counselling for Depression: a Comparison of therapies and Relationship to Outcome.

Researcher's name: Liz Gilley

The recording you are coding has been collected as part of a research project. Recordings may contain information of a very personal nature, which should be kept confidential and not disclosed to others. Maintaining this confidentiality is of utmost importance to the University.

We would like you to agree:

1. Not to disclose any information you may hear on the recording to others,
2. Only to accept files provided on an encrypted memory stick
3. Only to use encrypted computers to listen to therapy recordings.
4. To keep the rating forms in a secure locked place when not in use and write the client identifying code on it, rather than the client name.
5. When coding a recording ensure it cannot be heard by other people,
6. To adhere to the Guidelines for coders (appended to this document) in relation to the use of computers and encrypted digital recorders, and
7. To show your ratings only to the relevant individual who is involved in the research project.
8. If you find that anyone speaking on a recording is known to you, we would like you to stop rating work on that recording immediately and inform the person who has commissioned the work.

Appendix O: Rater Confidentiality Form and Guidelines continued

Declaration

I have read the above information, as well as the Guidelines for Coders, and I understand that:

1. I will discuss the contents of the recording only with the individual involved in the research project
2. I will only access files provided via secure memory stick on an encrypted computer and not reveal this login to anyone.
3. I will keep the rating forms in a secure place when not in use
4. When coding a recording I will ensure it cannot be heard by others
5. I will treat the coding of the recording as confidential information
6. I will adhere to the requirements detailed in the Guidelines for raters in relation to coding recordings.
7. If either person on the recordings is known to me I will undertake no further work on the recording

I agree to act according to the above constraints

Your name _____

Signature _____

Date _____

Occasionally, the conversations on recordings can be distressing to hear. If you should find it upsetting, please stop the coding and raise this with the researcher as soon as possible.

Appendix O: Rater Confidentiality Form and Guidelines continued

Guidelines for coders

Introduction

The course has created the guidelines below for anyone who is involved in coding data for staff or trainees in the Clinical Psychology Unit, University of Sheffield.

In addition to adhering to the following guidelines, **coders must sign a confidentiality form** prior to beginning any work. If you are unsure about any of the information given below, or for a copy of the confidentiality form, please contact the relevant trainee/member of staff.

When undertaking coding, whether from tapes or digital recording, you must:

- Password protect the computer files you are typing **before you type any text** – this can be done easily in Microsoft Word (instructions below)

- Anonymise any personal information contained in the data you are transcribing as you type e.g. names. Please contact trainee or member of staff who transcribing you are doing if you have any queries about this.

- Delete any files from your computer (including from your 'Trash' folder) once you have submitted your completed ratings.

- Keep the rating forms in a secure locked place when not in use.

Appendix O: Rater Confidentiality Form and Guidelines continued

Instructions for a password protecting files on a PC:

For Word 1998-2003:

1) Open a blank Word document
2) Go to Tools on the menu bar and select Options
3) Go to the Security tab and insert a password to open the document. You will be asked to re-type this, then please ensure you click ok before closing the Options menu.

For Word 2007:

1) Open a blank Word document
2) Go to Save As and choose the compatible mode
3) Click Tools, then select General Options
4) Enter a password to open the document. You will asked to re-type this, then please ensure you click ok before closing the dialogue box.

Instructions for password protecting files on a Mac:

1) Open a blank Word document
2) Go to Word on the menu bar and select Preferences
3) Click on Security and insert a password to open the document. You will be asked to re-type this, then click ok.

Appendix P: Double Rated Coding Form

Double rated therapy recording discussion sheet

CODER INITIALS (both):

DATES OF CODING (both):

DATE OF THERAPY SESSION:

CLIENT NUMBER:

THERAPIST NUMBER:

SESSION NUMBER:

DATE OF DISCUSSON:

ITEMS RATED DIFFERENTLY (please list):

ITEMS RATED WITH DIFFERENCE GREATER THAN 1 (please list):

Summary of discussion of differences: