University of Cumbria

Cameron, Sarah, Rodgers, Jacqui and Dagnan, David ORCID: https://orcid.org/0000-0002-5709-1586 (2018) The relationship between the therapeutic alliance and clinical outcomes in cognitive behavior therapy for adults with depression: a meta-analytic review. Clinical Psychology & Psychotherapy, 25 (3). pp. 446-456.

Downloaded from: http://insight.cumbria.ac.uk/id/eprint/8369/

Usage of any items from the University of Cumbria's institutional repository 'Insight' must conform to the following fair usage guidelines.

Any item and its associated metadata held in the University of Cumbria's institutional repository Insight (unless stated otherwise on the metadata record) may be copied, displayed or performed, and stored in line with the JISC fair dealing guidelines (available here) for educational and not-for-profit activities

provided that

- the authors, title and full bibliographic details of the item are cited clearly when any part of the work is referred to verbally or in the written form
- a hyperlink/URL to the original Insight record of that item is included in any citations of the work
- the content is not changed in any way
- all files required for usage of the item are kept together with the main item file.

You may not

- sell any part of an item
- refer to any part of an item without citation
- amend any item or contextualise it in a way that will impugn the creator's reputation
- remove or alter the copyright statement on an item.

The full policy can be found here.
Alternatively contact the University of Cumbria Repository Editor by emailing insight@cumbria.ac.uk.

The relationship between the therapeutic alliance and clinical outcomes in Cognitive Behaviour Therapy for adults with depression: A meta-analytic review

Authors

Sarah Kate Cameron[1], Jacqui Rodgers[2], Dave Dagnan[3].

NHS Lanarkshire[1]

Newcastle University[2]

Cumbria Partnership NHS Foundation Trust[3]

sarah.cameron@lanarkshire.scot.nhs.uk

Keywords: alliance, therapeutic alliance, depression, cognitive therapy, cognitive behaviour therapy

Abstract

Research consistently provides evidence for the relationship between the therapeutic alliance (TA) and outcome across various therapies and presenting problems. Depression is considered the leading cause of disability worldwide and there is substantial evidence for the efficacy for Cognitive Behaviour Therapy (CBT) in its treatment. At present there is lack of clarity specifically about the relationship between the TA and outcome in CBT for depression. The present review is the first meta-analytic review to explore this relationship whilst also considering moderators. Within a random-effects model, an overall mean effect size of $r = 0.26$ (95% CI .19 - .32) was found indicating that the TA was moderately related to outcome in CBT for depression. The mean TA-outcome correlation is consistent with existing meta-analysis that looked across a broad range of presenting problems and psychological therapies. A secondary exploratory analysis of moderators suggested the TA-outcome relationship varied according to the TA rater, where the relationship was weaker for therapist raters compared to clients and observer raters. Additionally, the results indicated that the TA-outcome relationship marginally increased over the course of CBT treatment. The results of the meta-analysis are discussed in reference to the wider body of research, methodological limitations, clinical implications and future directions for research.

Key practitioner message

1. Development of the therapeutic alliance may promote clinical outcomes in CBT for depression. Practitioners would do well to employ efforts on fostering the therapeutic alliance with clients.
2. Measurement and monitoring of the therapeutic alliance is recommended, allowing any difficulties in the alliance to be identified and addressed by the therapist.
3. The therapeutic alliance is related to outcome at all stages of therapy, although it may be strongest at late stages of therapy. A focus on maintaining the therapeutic alliance at all stages would be important, including early phases.
4. Therapists and clients may hold different views of their therapeutic alliance. Monitoring both client and therapist's views of the alliance would be particularly helpful.

Introduction

Depression is deemed as the leading cause of disability and is there are an estimated 350 million people effected by depression worldwide (WHO, 2015). Major Depressive Disorder (MDD) is marked by persistent depressed mood and diminished interest and pleasure in activities *(DSM – 5; APA, 2013)*. Cognitive Behaviour Therapy (CBT) for depression has been described as "structured, short-term, present-orientated psychotherapy for depression, directed toward solving problems and modifying dysfunctional thinking and behaviour" (as cited in J.S. Beck, 1995 p. 1). CBT for depression has a strong evidence-base and is deemed to be highly efficacious (e.g. Gloaguen, 1998; Tolin, 2010). There is much research that has investigated the TA within psychological therapy for depression.

The therapeutic alliance (TA), therapeutic relationship or working alliance are terms used to describe related constructs (Norcross, 2002), which are broadly defined as the 'collaborative and affective bond between therapist and client' (Martin, Garske and Davis, 2000). The TA is viewed as a factor common to all treatments that is integral in determining positive therapy outcomes, at times considered as more significant than the treatment itself (Messer & Wampold, 2002). The TA has been extensively investigated in general adult samples and a strong TA has been consistently proposed to have a relationship with positive outcomes across various psychological therapies, with moderate effect sizes found in meta-analyses (e.g. Martin *et al*, 2000, $r = .22$, $k = 79$; Horvath *et al*, 2011, $r = .275$, $k = 190$).

There is no one agreed TA definition and it is a concept with a complex history. The concept of the TA can be traced back to Freud (1912) and psychoanalysis were he discussed that the client held a therapist 'within them' which supported therapeutic progress. Freud proposed both collaboration between client and therapist in their desire to address the client's 'pain' and the concepts of transference and counter-transference as the core change processes in therapy (Elvins and Green, 2008).

Zetzel (1956) was the first to use the term 'therapeutic alliance' and in line with a psychoanalytic approach, suggested the client used the healthy part of their 'ego' to connect with the therapist, which enabled engagement in the therapeutic tasks. Subsequently Greenson (1965) suggested the 'therapeutic alliance' was the bond between client and therapist, and the 'working alliance' was the client's ability to engage with tasks. Orlinsky and Howard (1975) went on to propose three dimensions that formed the TA; working alliance (investment by both therapist and client in the therapy), empathic resonance and mutual affirmation.

In 1979, Bordin reviewed the TA literature and the psychodynamic origins of the TA, moving to propose a pan- theoretical model. Bordin (1979) conceptualised the 'working alliance' and proposed it comprised three elements; agreement of goals, assignment of tasks and development of bond and in combination these factors determine the strength and quality of the alliance.

Although the TA was more traditionally associated with psychodynamic therapies, it has been increasingly acknowledged as a key process within Cognitive Behaviour Therapy (CBT; Raue, 1997). Indeed, Beck (1995) emphasized that the TA is one of six basic principles of CBT and facilitates change. It is considered necessary to promote change, but not sufficient without the actual therapeutic techniques (Beck, 1995).

There is research that suggests that the strength and nature of TA differs across therapy types (Barber *et al*, 2006). For example, Raue (1997) found the TA was significantly higher in CBT compared to psychodynamic psychotherapy. However, meta-analyses have suggested that there are no differences in the strength of the TA-outcome relationship across therapies, e.g. Horvath *et al* (2011) looked at the alliance-outcome relationship across psychological therapies and problems. CBT was examined as a subgroup and reported an effect size of $r = .35$, and there were no significant difference across different therapies (e.g. psychodynamic, interpersonal therapies).

Some have suggested that the TA is not related to outcome in CBT. Ilardi and Craighead (1994) conducted a narrative review around the importance of non-specific factors in CBT for depression and a small aspect of this review looked at the TA and outcome. They reviewed three studies and suggested that there was little evidence that the TA in early stages of therapy influenced subsequent outcome. Whereas, one component of further narrative review by Keijsers, Schaap and Hoogduin (2000), which explored patient and therapist behaviour more generally and its influence on outcome, reported the association between the TA and outcome in CBT for anxiety and depression. This review proposed there was evidence of relationship between TA in CBT and outcome for depression.

Previous meta-analyses have provided effect sizes for the relationship between alliance and outcome in CBT across different problems. There is strong evidence for the relationship between TA and outcome across various therapies and presenting problems. Although CBT contains common elements across problems, there are disorder-specific models and the components of CBT for different problems can have differing emphasis. Additionally, there are key differences between psychological therapies which could influence each therapy's association with the TA.

The TA-outcome relationship is likely to be complex. There has been debate in the CBT literature whether it is the TA that leads to better outcome, or symptom change results in a stronger TA. DeRubeis and Feeley (1990) attempted to understand these two relationships further specifically within CBT with participants with depressive disorders**.** They found that prior depressive symptom improvement was related to a stronger TA, whereas they did not find a significant relationship between the TA and subsequent depression change. Whereas, Webb *et al* (2011) found that the TA predicated greater depressive symptom reduction in participants with depression engaged in CBT, controlling for 'temporal comfounds' (prior depressive symptom

change). In addition, Webb *et al* suggested that the TA assessed at late sessions was predicted by prior symptom change.

This review specifically focuses on CBT for depression, which has a good evidence base, many studies in this area and common use in clinical practice. It is pertinent, therefore, to conduct a closer examination of the evidence for the relationship between TA and outcome in CBT for depression. The most recent review in this area was some time ago, in year 2000, and therefore an update is pertinent. Moreover, there is no existing meta-analysis in this area.

The present review aims to use all currently available data to establish the strength of the relationship between the TA and outcome in CBT for depression, whilst also considering moderators. The main objective is to establish the mean TA-outcome correlation in CBT for depression. A secondary objective is to understand the influence of two moderators of the TA-outcome relationship: time of TA measure and rater of TA.

Method

The research papers included in the present meta-analysis were selected following a rigorous search and screening process. Search terms and selection criteria were developed by the authors to identify the literature relevant to the review question. They were utilised in an initial pilot review, which had the same question. The pilot review indicated that the search terms were broad and resulted in a wide range of studies, including those that were relevant to the research question.

Search terms

The search terms were selected by identifying key words used within the field. These key words were identified by consulting the literature and diagnostic guidelines; DSM-III – R, DSM – IV and DSM – 5. Wild cards (*) were used to ensure that all variations of the search terms were captured.

Therapeutic Alliance: alliance OR therapeutic alliance OR working alliance OR helping alliance OR therapeutic relationship OR therapeutic bond.

Cognitive Behavioural Therapy: cognitive behavi* (cognitive behaviour/al therapy, cognitive behavior/al therapy, cognitive behaviour/al therapies, cognitive behavior/al therapies) OR cognitive therap* (cognitive therapy, cognitive therapies).

Depression: Major depress* (Major depression, major depressive disorder, major depressive episode) OR depress* (depression, depressive disorder, depressive episode) OR dysthymi* (dysthymia, dysthymic disorder) OR persistent depress* (persistent depression, persistent depressive disorder, persistent depressive episode) OR depressi* neurosis (depression neurosis, depressive neurosis) OR neurotic depress* (neurotic depressive, neurotic depression).

Selection criteria

The selection criteria were: correlational research that reported a TA-outcome relationship, papers reporting primary research, details on the measurement of the TA and the method of diagnosing depression, a CBT intervention delivered with the cognitive model underlying the intervention. Interventions with overlaps with CBT, for example, behavioural activation or relaxation were excluded as the underlying model is different and there is absence of the 'cognitive' component. CBT interventions delivered on an individual, face to face basis were included (group interventions would be excluded). Further selection criteria were that depression was the main target or the intervention and depressive symptoms as the primary outcome variable. Finally, the focus of the review was on an adult population, therefore participants in the studies were 18 years and above. No studies were excluded based on language and no date limits were set.

Selection of studies

The first author (SC) took the lead in completing the searches and screening the results, based on the selection criteria. Firstly, a comprehensive electronic database search was conducted using the key search terms. The following electronic databases were searched; MEDLINE(R) 1946 to December 2015 (via Ovid), Embase (Ovid SP) 1974 to December 2015, PsycINFO (Ovid SP) 1806 to December 2015 and Scopus (to December 2015). A total of 1,026 results were generated after removing duplicates. The main reason for exclusion was that depression was not the primary outcome measure, for example days to relapse or 'maladjustment.' Studies were also excluded due to having an inappropriate sample (e.g. depression related to a physical health or a mixed anxiety and depression sample), or CBT not being delivered on an individual, face to face format. Further reasons for exclusion were absence of data to answer the review question, for example TA and depression were measured but the relationship between these variables were not reported or the analysis was conducted across CBT and other treatment conditions. After screening these papers according to the selection criteria, 11 papers were identified. Authors were contacted for additional data for five papers, where the TA-outcome correlation was not available in the paper, with three authors providing this data.

In addition to an electronic database search, a series of further searches were also completed. These included a search of identified prominent authors in the field, a search of key journals, a hand search of reference sections of key papers and a search of unpublished papers. The author search involved entering the second names of the authors into the Scopus data, with no date range. The following results were generated; Webb (228 results), Blatt (196 results), DeRubeis (204 results), Zuroff (152 results), Hardy (98 results) and Strunk (39 results). After screening of the results with reference to the selection criteria no additional papers were identified. Either the paper did not meet selection criteria or papers that had already been identified in the electronic database.

Subsequently, searches of three key journals, containing many of the published studies, using the terms 'alliance CBT depression,' and no date range indicated, yielded 553 results in total; Journal of Consulting and Clinical Psychology (434 results), Behaviour Therapy and Research (95 results) and Cognitive Therapy and Research (24 results). Hand searches of the reference sections of the 11 papers from the electronic database search and three key meta-analyses journal searches (Horvath *et al*, 2011, Martin *et al*, 2000 and Flückiger *et al*, 2012) were conducted. To increase chances to identifying any unpublished data a ProQuest database search was conducted, which resulted in 256 results. To aim to capture the most up to date papers, a Google Scholar search using the search terms 'alliance CBT depression and date range from 2014 to December 2015, resulted in 2,000 results. The series of searches described were reviewed with reference to the selection criteria and resulted in 1 additional paper, which was from the ProQuest database search.

Following the searches described, there were a total of 15 papers. However, two articles were subsequently excluded as they were reporting on shared datasets with other studies. The authors decided that when more than 1 paper reported the same data, the first published papers were selected. Thus, 13 papers were included in the meta-analysis. Further information on study selection is provided below in *Figure 1*.

Insert Figure 1 about here

Meta-Analytic Method

Thirteen papers were included in the meta-analysis. Each study was coded by the lead researcher using the following information: author name, number of participants, details of TA measure, rater of the TA (client/self-report, therapist and observer rated), time of TA measure (early, mid and late), details of outcome measures and *r*-type effect size. To check the accuracy

of coding, a second-rater coded a sample of the final papers (six papers) and there was 100% agreement between the raters.

The relationship between TA and therapy outcome in the final studies was calculated in three ways (1) partial correlation between TA and post-treatment depression, controlling for pre-treatment depression, (2) partial correlation between TA and residualised change in depression scores over treatment, controlling for the previous session depression scores and (3) correlation between TA and residualised change in depression scores over treatment. Due to the differences in how the outcomes were presented there were both negative and positive relationships. The decision was made by the authors to code all the effect sizes as a positive correlation, so they were comparable.

Independent effect sizes

One independent effect size for the relationship between TA-outcome for each study was included in the meta-analysis. A number of studies reported multiple effect sizes as they used multiple measures or raters for each construct, therefore these effect sizes were not independent of each other. For the papers when more than one effect size was reported, a mean effect size was calculated for each study across measures of each construct (TA and depression), TA rater and time of TA measure. Whether the effect sizes used were means or from single reported associations are shown in *Table 1*.

For three studies (Stiles-Shields et al, 2014; Teismann et al, 2012; Webb et al, 2014) there was drop out over the course of treatment and therefore each time point had a different sample size. For these studies, the effect size for the early point of treatment was extracted, rather than generating a mean effect size.

Meta-analysis software

Based on the procedures outlined, for the main analyses 13 independent *r*-type effect sizes were entered into the Comprehensive Meta-Analysis software (Version 3.0; Borenstein et al, 2014).

Homogeneity and Heterogeneity analysis

In meta-analysis, it is important to investigate the homogeneity of effect size distributions; whether the effect sizes that are averaged into a mean value estimate the same population effect size (Hedges, 1982). There are two primary statistical indicators of homogeneity, Q and I^2 *s*tatistics.

The random effects-model was chosen for this study as it assumes the true effect size may vary between studies and is more appropriate for studies which have sources of variance. The random effects model was chosen due to there being variability within the research area, for example, various measures, potential differences in the delivery of CBT and differences between therapists.

Outliers

Outlier effect sizes were also considered to investigate their potential impact on the estimation of the strength of the TA-outcome relationship. The sample-adjusted meta-analytic deviation (SAMD) statistic was calculated to test for the presence of statistical outliers (Huffcutt and Arthur, 1995).

Moderator analysis

Time of TA measure and the rater of the TA measure may impact on the TA-outcome relationship, therefore it was important to explore this further. A secondary analysis was conducted looking at disaggregated effect sizes as a way of exploring the influence of TA rater and time point as moderators. Effect sizes were weighted based on the proportion of the original sample sizes. Although this breaks the assumptions of non-independent effect sizes, it is

considered a trade-off to allow exploration of these factors. Differences between groupings due to moderators were not examined for significance because the effect sizes were not independent.

Results

Thirteen studies were included in the investigation of the relationship between TA and outcome in CBT for depression. *Table 1* presents the details of the final papers for the review. These studies were conducted over 26 years between 1989 and 2015. The sample sizes varied from $n = 21$ and $n = 128$ and the effect sizes for the correlations between TA and outcome ranged from $r = 0.06$ to $r = 0.59$. *Figure 2* details the Forest Plot, where each horizontal line represents the study correlation and the 95% confidence interval, the black diamond represents the overall correlation of effect sizes and the black squares represents the weight of the effect size.

The overall mean correlation between TA and outcome was $r = 0.26$ (95% CI .19 - .32). The effect sizes were homogeneous ($Q = 9.66$, df 13, $p = 0.646$), suggesting that the effect sizes were measuring a single population of effects. The I^2 statistic was computed and suggested that heterogeneity was low within the group of effect sizes overall ($I^2 = 0\%$).

Insert Table 1 about here

Insert Figure 2 about here

Publication bias

Publication bias recognises that studies with significant results have a higher chance of publication, whereas non-significant results are less frequently published. A funnel plot was produced to examine publication bias (see *Figure 3*), where the clear dots are the effect sizes included in the meta-analysis. In the presence of publication bias, there would be a high concentration of studies to the left of the mean appearing towards the bottom of the graph. The studies are distributed almost symmetrically at the top of the graph, suggesting absence of publication bias.

Orwin's failsafe N (Orwin, 1983) was used also to account for publication bias, which identified that 155 studies that accepted the null hypothesis would be required to reduce the correlation to a trivial value, i.e. $r = <0.01$.

Insert Figure 3 about here

Moderator analysis

Moderator analyses was not be warranted as the effect sizes were homogeneous, however on inspection of the effect sizes in the final papers, it appeared that averaging effect sizes across TA time point and rater was reducing the variance in the TA-outcome correlation. A secondary exploratory analysis examined TA time point and rater as moderators of the TA-outcome relationship. To examine these moderators, mean effect sizes were computed for TA rater and time point of TA measurement. *Table 2* details 23 effect sizes for time point and 18 effect sizes for rater, extracted from the 13 final papers. Significance between groupings within the moderators were not examined due to the effect sizes being non-independent of each other.

Time point of therapeutic alliance measure

The number of effect sizes for the TA-outcome relationship according to the time point of TA measurement in treatment was as follows; early ($n = 8$), mid ($n = 5$), late ($n = 6$) and average ($n = 8$). The mean effect sizes for the relationship between TA and outcome measured at early, mid and late in treatment were $r = 0.21$ (95% CI .12 - .31); $r = 0.28$ (95% CI .11 –.43) and $r = 0.27$ (95% CI .13 – .40). When the TA was measured at various time points and an average was presented in the study, the relationship with outcome was $r = 0.42$ (95% CI .22 - .59). These results suggest that the strength of the TA-outcome relationship when measured mid and late in treatment is a marginally higher than early ratings. Although for the 'average' time point the TA-outcome was strongest, there was a higher level of heterogeneity $I^2 = 70.5\%$, therefore this finding should be interpreted with more caution.

Rater of therapeutic alliance

Client ratings of the TA represented the largest number of effect sizes for the TA-outcome relationships (n = 9). There was a smaller, but equal number for the other raters of the TA; therapist (n = 4) and observer (n = 4). There was only one effect size for 'average', so was excluded from this analysis. The effect sizes according to TA rater were: client (r = 0.27, 95% CI .19 – .36), therapist (r = 0.19, 95% CI .03 - .33) and observer (r = 0.41, 95% CI .22 – .58). The TA-outcome relationship for therapist ratings was weaker, when compared with client and observer ratings. Although, again, there was a higher level of heterogeneity for the observer effect size (I^2 = 54.7%), therefore this finding should be interpreted with more caution.

Insert Table 2 about here

Discussion

The primary aim of the meta-analysis was to examine the available literature on the TA-outcome relationship in CBT for depression. Following a systematic review, 13 papers were included in the main meta-analysis, resulted in an overall mean effect size of r = 0.26 (95% CI .19 - .32). This suggests that the TA is moderately related to outcome in CBT for depression. This finding is robust as Orwin's fail safe N statistic identified that approximately 155 studies that accepted the null hypothesis would be needed to reduce the correlation to a trivial value (r = <0.01). The effect sizes in the main analysis were homogeneous, suggesting that the correlation represents a single population of effects. The mean TA-outcome correlation is consistent with existing meta-analyses that looked across various presenting problems and psychological therapies (e.g. Martin *et al*, 2000, r = .22, k = 79; Horvath *et al*, 2011, r = .275, k = 190). Previous meta-analyses by Horvath *et al* (2001, 2011) found a high level of heterogeneity across studies, which is very different to this review. This may be due to the present meta-analysis focusing on a specific presenting problem (depression) and therapy (CBT).

Although this finding may appear modest, it could be considered large in comparison to the relationship between other process variables and outcome in CBT. For example, Webb *et al* (2010) reported mean effect sizes of the relationships between therapist adherence to treatment protocol and therapist competence with therapy outcome as $r = 0.02$ and $r = 0.07$ respectively.

In this meta-analytic review, a secondary exploratory analysis examined moderators; TA time point and rater. These results suggest that the strength of the TA-outcome relationship when measured mid and late in treatment was marginally higher than early treatment ratings. This suggestion is made tentatively due to the small increase in the TA-outcome relationship later in treatment and because significant differences were not examined. Nevertheless, the finding that the TA-outcome association was stronger later in treatment may not be surprising and is consistent with previous meta-analysis (Horvath *et al*, 2011). It is also interesting to consider whether this finding is influenced by prior depression change, in that a stronger TA is related to greater symptom improvement as therapy progresses. There has been some debate in the CBT literature with regards to whether it is the TA that leads to better outcome, or symptom change results in a better quality TA (e.g. De Rubeis and Feeley, 1990). This is a complex issue and further research is needed, it may not be a linear process but cyclical, i.e. TA leads to better outcome, which leads to a stronger TA and so on.

For TA rater, the TA-outcome relationship was weaker for therapist ratings, when compared with client and observer ratings. Although the difference for this moderator was larger, it is emphasised that it was not possible to examine whether there were significant differences between the TA-outcome relationships according to the rater. The findings in this review are similar to previous research that suggests it is both observer and self-report rated TA measures are the superior predictors of therapy outcomes compared to therapist-rated TA measures (Horvath and Symonds, 1991). In considering the differences between the TA-outcome relationships according to the rater, it may be that therapist's knowledge and role in therapy is

different, perhaps they are viewing the TA in terms of their theoretical knowledge and clinical experience. Contrastingly, the clients are focusing more directly on their felt experience.

The analysis has various limitations. It was identified that the TA-outcome was principally calculated by partial correlation, either between TA and post-treatment depression, controlling for pre-treatment depression or between TA and residualised change in depression scores over treatment, controlling for the most recent session's depression score. Additionally, one study included a correlation between TA and subsequent depression change. The difficulties of using partial correlations in meta-analyses are acknowledged, including the potential influence on inflating the magnitude and altering the direction of the effect (Aloe, 2014). It is recommended that synthesis of partial correlations, is appropriate if the correlations are controlling for the same variables (Aloe, 2014). Synthesis of partial correlations in this meta-analysis is acknowledged as a limitation. However, this approach was important for this particular meta-analysis so that previous depression severity was controlled for when considering the TA-outcome relationship to allow comparison between participants of varying depression severity from the outset within a study.

In the moderator analyses the meta-analytic assumptions of independence were broken. These analyses were exploratory and significant differences between groupings were not investigated because the effect sizes were not independent. In addition, effect sizes from the same study were weighted appropriately to ensure their contribution was not overemphasised. While this is a limitation of the current review, it was considered a trade-off to allow for exploration of important factors which may influence the TA-outcome relationship. It appears that the rater has some influence on the TA-outcome relationship. It should also be noted that for the TA rater analysis, two of the subgroups were small and it was unbalanced, where *n* represent the number of effect sizes; Client ($n = 9$), Therapist ($n = 4$) and Observer ($n = 4$).

The TA is recognized as increasingly important in CBT although the focus remains on therapy technique. In terms of process variables, this study shows the importance of alliance to outcome in CBT. Therapists delivering CBT would do well to attend to building and maintaining the alliance in their work with patients. In terms of collecting outcomes to monitor change in therapy, due to the evidence of a TA-outcome relationship in CBT, it is recommended that therapists also routinely measure the TA. This would allow therapists to monitor the strength of the TA in different areas and, if there are weaknesses, consider how they can boost the TA further.

It is helpful to consider how the TA could be promoted in CBT. There has also been consideration around specifically targeting the improvement of the TA within CBT. Castonguay *et al* (2004) developed Integrated Cognitive Therapy (ICT) for depression, which incorporated 'alliance rupture repair techniques.' Promising results were reported in an initial study, with a large effect size (d =1.91) for pre-post BDI scores. When ICT was compared with standard CBT (n=11 per group), ICT had superior depression improvement on the BDI (d = 0.5) and higher TA ratings than standard CBT (Constantino *et al*, 2008). Further research is required in this area to further establish whether ICT is superior to CBT.

The present review focused on the relationship between TA and subsequent depression change in CBT for depression. There has been an ongoing debate about whether TA is a cause or consequence of outcome. Some have suggested that in fact prior depression change leads to a stronger TA, rather than the TA leading to better outcomes. Further research is required to further understand this complex relationship.

There are a range of factors which may influence the TA-outcome relationship, for example therapist experience, fidelity to CBT model, treatment duration and so on. It was beyond the scope of this review to examine these factors, however future meta-analysis could investigate their relevance in CBT for depression. Problem type and severity are patient factors which are

known to influence development of the TA (Horvath *et al*, 2011). There are known to be particular difficulties in personality disorder and substance misuse problems. Further meta-analysis could focus on the TA-outcome relationship for these problems.

In conclusion, this meta-analytic review provides evidence of a moderate relationship between the TA and outcome in CBT for depression, with an overall mean effect size of $r = 0.26$ (95% CI .19 - .32). A secondary exploratory analysis of moderators suggested that the TA-outcome relationship varied according to TA rater, where the relationship was weaker for therapists than clients and observer raters. Additionally, the results indicated that the TA-outcome relationship marginally increased over the course of CBT treatment. The results of this meta-analysis highlights the need for attention to be given to the TA when CBT is being delivered in clinical practice.

Key practitioner message

1. Development of the therapeutic alliance may promote clinical outcomes in CBT for depression. Practitioners would do well to employ efforts on fostering the therapeutic alliance with clients.
2. Measurement and monitoring of the therapeutic alliance is recommended, allowing any difficulties in the alliance to be identified and addressed by the therapist.
3. The therapeutic alliance is related to outcome at all stages of therapy, although it may be strongest at late stages of therapy. A focus on maintaining the therapeutic alliance at all stages would be important, including early phases.
4. Therapists and clients may hold different views of their therapeutic alliance. Monitoring both client and therapist's views of the alliance would be particularly helpful.

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*(2), 155-172.

Alexander, L. B., & Luborsky, L. (1986b). The Penn Helping Alliance Scales. In L. S. G. W. M. Pinsof (Ed.), *The psychotherapeutic process: A research handbook* (pp. 325-366). New York, NY, US: Guilford Press.

Aloe, A. M. (2014). An empirical investigation of partial effect sizes in meta-analysis of correlational data. *The Journal of general psychology, 141*(1), 47-64. doi:10.1080/00221309.2013.853021

APA. (1980). *DSM-III-R: Diagnostic and statistical manual of mental disorders*: American Psychiatric Association.

APA (2000). *Diagnostic and statistical manual of mental disorders: DSM-IV-TR®*: American Psychiatric Pub.

APA (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*: American Psychiatric Pub.

Andresen, E. M., Malmgren, J. A., Carter, W. B., & Patrick, D. L. (1994b). Screening for depression in well older adults: evaluation of a short form of the CES-D. *American journal of preventive medicine*.

Barrett-Lennard, G. T. (1962). Dimensions of therapist response as causal factors in therapeutic change. *Psychological Monographs: General and Applied, 76*(43), 1-36. doi:10.1037/h0093918

Beck, A. T., Steer, R. A., Ball, R., & Ranieri, W. F. (1996). Comparison of Beck Depression Inventories-IA and-II in psychiatric outpatients. *Journal of personality assessment, 67*(3), 588-597. doi:10.1207/s15327752jpa6703_13

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 4*(6), 561.

Beck, J. S. (1995). *Cognitive therapy: Basics and beyond*. New York, NY: Guilford Press.

Bordin, E. S. (1979). The generalizability of the psychoanalytic concept of the working alliance. *Psychotherapy: Theory, Research & Practice, 16*(3), 252. doi:10.1037/h0085885

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2014). Comprehensive meta-analysis version 3. *Englewood, NJ: Biostat*, *104*.

*Carter, J. D., Crowe, M. T., Jordan, J., McIntosh, V. V. W., Frampton, C., & Joyce, P. R. (2015). Predictors of response to CBT and IPT for depression; the contribution of therapy process. *Behaviour research and therapy, 74*, 72-79. doi:10.1016/j.brat.2015.09.003

*Castonguay, L. G., Goldfried, M. R., Wiser, S., Raue, P. J., & Hayes, A. M. (1996). Predicting the effect of cognitive therapy for depression: A study of unique and common factors. *Journal of Consulting and Clinical Psychology, 64*(3), 497-504. doi:10.1037/0022-006X.64.3.497

Castonguay, L. G., Schut, A. J., Aikens, D. E., Constantino, M. J., Laurenceau, J.-P., Bologh, L., & Burns, D. D. (2004). Integrative Cognitive Therapy for Depression: A Preliminary Investigation. *Journal of Psychotherapy Integration, 14*(1), 4-20. doi:10.1037/1053-0479.14.1.4

Constantino, M., & Smith-Hansen, L. (2008). Patient interpersonal factors and the therapeutic alliance in two treatments for bulimia nervosa. *Psychotherapy Research, 18*(6), 683-698. doi:10.1080/10503300802183702

*DeRubeis, R. J., & Feeley, M. (1990). Determinants of change in cognitive therapy for depression. *Cognitive Therapy and Research, 14*(5), 469-482.

Elvins, R., & Green, J. (2008). The conceptualization and measurement of therapeutic alliance: An empirical review. *Clinical Psychology Review, 28*(7), 1167-1187. doi:doi:10.1016/j.cpr.2008.04.002

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*(1), 10. doi:10.1037/a0025749

Freud, S. (1912). *The dynamics of transference*. New York, NY: W. A. White and S. E. Jelliffe.

Gaston, L. (1991). Reliability and criterion-related validity of the California Psychotherapy Alliance Scales—patient version. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 3*(1), 68. doi:10.1037/1040-3590.3.1.68

Gloaguen, V., Cottraux, J., Cucherat, M., & Ivy-Marie, B. (1998). A meta-analysis of the effects of cognitive therapy in depressed patients. *Journal of Affective Disorders, 49*(1), 59-72. doi:http://dx.doi.org/10.1016/S0165-0327(97)00199-7

Greenson, R. R. (1965). The working alliance and the transference neurosis. *The Psychoanalytic Quarterly, 34*(2), 155-179.

Hamilton, M. (1960). A rating scale for depression. *Journal of neurology, neurosurgery, and psychiatry, 23*(1), 56.

Hatcher, R. L., & Gillaspy, J. A. (2006). Development and validation of a revised short version of the Working Alliance Inventory. *Psychotherapy Research, 16*(1), 12-25.

Hedges, V. (1982). Fitting categorical models to effect sizes from a series of experiments. *7*, 119-137. doi:10.2307/1164961

Horvath, A. O. (2001). The alliance. *Psychotherapy: Theory, research, practice, training, 38*(4), 365-372. doi:10.1037/0033-3204.38.4.365

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

Horvath, A. O., Del Re, A. C., Flückiger, C., & Symonds, D. (2011). Alliance in individual psychotherapy. *Psychotherapy, 48*(1), 9-16. doi:10.1037/a0022186

Horvath, A. O., & Symonds, B. D. (1991). Relation between working alliance and outcome in psychotherapy: A meta-analysis. *Journal of counseling psychology, 38*(2), 139-149. doi:10.1037/0022-0167.38.2.139

Huffcutt, A. I. A., Winfred (1995). Development of a new outlier statistic for meta-analytic data. *80*, 327-334. doi:10.1037/0021-9010.80.2.327

Ilardi, S. S., & Craighead, W. E. (1994). The Role of Nonspecific Factors in Cognitive-Behavior Therapy for Depression. *Clinical Psychology: Science and Practice, 1*(2), 138-155. doi:10.1111/j.1468-2850.1994.tb00016.x

Keijsers, G. P. J., Schaap, C., & Hoogduin, C. A. L. (2000). The Impact of Interpersonal Patient and Therapist Behavior on Outcome in Cognitive-Behavior Therapy A Review of Empirical Studies. *Behavior Modification, 24*(2), 264-297.

*Löffler, U., Sakalli, B., Stopsack, M., Mander, J., Bents, H., & Barnow, S. (2014). Relationships between patients' ratings of alliance and therapists' ratings of CBT techniques: A naturalistic study in depressive disorders. *Verhaltenstherapie und Verhaltensmedizin, 35*(4), 334-350.

*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*(2), 368-373. doi:10.1037/a0035994

Luborsky, L., Barber, J. P., Siqueland, L., Johnson, S., Najavits, L. M., Frank, A., & Daley, D. (1996). The revised Helping Alliance questionnaire (HAq-II): psychometric properties. *The Journal of psychotherapy practice and research, 5*(3), 260.

*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*(8), 464-472.

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*(3), 438. doi:10.1037/0022-006X.68.3.438

Messer, S. B., & Wampold, B. E. (2002). Let's Face Facts: Common Factors Are More Potent Than Specific Therapy Ingredients. *Clinical Psychology: Science and Practice, 9*(1), 21-25. doi:10.1093/clipsy.9.1.21

Montgomery, S. A., & Asberg, M. (1979). A new depression scale designed to be sensitive to change. *The British Journal of Psychiatry, 134*(4), 382-389. doi:10.1192

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

O'Malley, S. S., Suh, C. S., & Strupp, H. H. (1983). The Vanderbilt Psychotherapy Process Scale: A report on the scale development and a process-outcome study. *Journal of Consulting and Clinical Psychology, 51*(4), 581. doi:10.1037/0022-006X.51.4.581

Orlinsky, D. E., & Howard, K. I. (1986). *The psychological interior of psychotherapy: Explorations with the Therapy Session Reports*. New York, NY, US: Guilford Press.

Orwin, R. G. (1983). A fail-safe N for effect size in meta-analysis. *Journal of educational statistics*, 157-159. doi:10.2307/1164923

*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*(4), 582-587. doi:10.1037/0022-006X.65.4.582

*Stiles-Shields, C., Kwasny, M. J., Cai, X., & Mohr, D. C. (2014). Therapeutic alliance in face-to-face and telephone-administered cognitive behavioral therapy. *Journal of Consulting and Clinical Psychology, 82*(2), 349-354. doi:10.1037/a0035554

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*(5), 730-738. doi:10.1037/a0029281

*Teismann, T., Michalak, J., Willutzki, U., & Schulte, D. (2012). Influence of rumination and distraction on the therapeutic process in cognitive-behavioral therapy for depression. *Cognitive Therapy and Research, 36*(1), 15-24. doi:10.1007/s10608-010-9347-6

Tolin, D. F. (2010). Is cognitive–behavioral therapy more effective than other therapies?: A meta-analytic review. *Clinical Psychology Review, 30*(6), 710-720. doi:10.1016/j.cpr.2010.05.003

Tracey, T. J., & Kokotovic, A. M. (1989). Factor structure of the Working Alliance Inventory. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 1*(3), 207. doi:10.1037/1040-3590.1.3.207

*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*(2), 143-157.

*Webb, C. A., Beard, C., Auerbach, R. P., Menninger, E., & Bjorgvinsson, T. (2014). The therapeutic alliance in a naturalistic psychiatric setting: Temporal relations with depressive symptom change. *Behaviour Research and Therapy, 61*, 70-77. doi:10.1016/j.brat.2014.07.015

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*(3), 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*(2), 200. doi:10.1037/a0018912

*Weck, F., Rudari, V., Hilling, C., Hautzinger, M., Heidenreich, T., Schermelleh-Engel, K., & Stangier, U. (2013). Relapses in recurrent depression 1 year after maintenance cognitive-behavioral therapy: The role of therapist adherence, competence, and the therapeutic alliance. *Psychiatry Research, 210*(1), 140-145. doi:10.1016/j.psychres.2013.05.036

WHO (1992). *The ICD-10 classification of mental and behavioural disorders: clinical descriptions and diagnostic guidelines*: Geneva: World Health Organization.

Zetzel, E. R. (1956). Current Concepts of Transference. *The International Journal of Psycho-Analysis, 37*, 369.

*Zuroff, D., Koestner, R., Moskowitz, D. S., McBride, C., Marshall, M., & Bagby, M. (2007). Autonomous motivation for therapy: A new common factor in brief treatments for depression. *Psychotherapy Research, 17*(2), 137-148. doi:10.1080/10503300600919380

Table 1: Independent effect sizes

Author and Date	Sample, Design & Country of Origin	*n* (CBT Sample)	Measure of Therapeutic Alliance	Rater of Therapeutic Alliance	Time of Therapeutic Alliance	Measure of Outcome	*r*
CARTER ET AL (2015)	Clinical sample Clinical Trial New Zealand	80	VTAS – R	O	E, M, L	MADRS	Average = 0.13
CASTONGUAY ET AL (1996)	Clinical Sample RCT (Hollon et al, 1992) USA	30	WAI – O	O	M, L	BDI HDRS	Average = 0.42
DERUBEIS AND FEELEY (1990)	Clinical sample Outpatient clinic USA	25	HAq	O	E, M, L	BDI	Average = 0.16
LÖFFLER ET AL (2014)	Clinical Sample Naturalistic study Germany	113	WAI-SR	C	E, L	BDI-II	Average = 0.36
LORENZO-LUACES ET AL (2014)	Clinical sample RCT (DeRubeis et al, 2005) USA	60	WAI – S – O	O	E	BDI-II	Early = 0.23

MARMAR ET AL (1989)	Clinical sample Older adult Clinical Trial	22	CALPAS	C, T	E, M, L	BDI	Early = 0.17
PRESCHL, MAERCHER AND WAGNER (2011)	Clinical sample RCT (online vs face to face CBT) Germany	28	WAI – S – C WAI – S – T	C, T	M, L	BDI – II (German version)	Average = 0.37
STILES-SHIELDS ET AL (2014)	Clinical sample RCT (online vs face to face CBT) USA	Early = 128 Late=121	WAI – SR WAI – SR – T	C, T	E, L	HRSD	Early = 0.28
TEISMANN ET AL (2012)	Clinical sample RCT Germany	Early = 67 Mid = 64 Late = 58	TEQ	C	E, M, L	BDI (German version)	Early = 0.29
TREPKA ET AL (2004)	Clinical sample Outpatient clinic (Data from Cahill et al, 2003) UK	21	CALPAS ARM	C	E, M, L	BDI BDI–II	Average = 0.59
WEBB ET AL (2014)	Clinical Sample Naturalistic setting (each patient has combined 1:1 and group) USA	103	WAI – S – C	C	E, M, L	CES-D	Early = 0.21
WECK ET AL (2013)	Clinical Sample RCT	80	HAq	C, T	A	HRSD – 15	Average = 0.2

	Germany						
ZUROFF ET AL (2007)	Clinical Sample RCT (McBride et al, 2006) Canada	36	CALPAS	C	E	HRSD BDI	Early = 0.06

Note.
Rater of Therapeutic Alliance: C = Client/Self - report, T = Therapist, O = Observer

Time of Therapeutic Alliance Measurement: E = Early, M = Mid, L = Late, A = Average.

Therapeutic Alliance Measures: Penn Helping Alliance Rating Scale (HAq; Morgan et al, 1982; Alexander and Luborsky, 1986); Barrett Lennard –Relationship Inventory (BL-RI; Barrett-Lennard, 1962); California Psychotherapy Alliance Scale (CALPAS; Marmar and Gaston, 1988); Agnew Relationship Measure – Version 3 (ARM; Agnew-Davies et al, 1998); Therapist Evaluation Questionnaire (TEQ; Schulte and Eifert 2002); Vanderbilt Therapeutic Alliance Scale Revised (VTAS-R; Krupnick, et al, 1996); Working Alliance Inventory (WAI; Horvath and Greenberg, 1989); Working Alliance Inventory – Short Form Client version (WAI-S-C; Tracy and Kokotovic, 1989); Working Alliance Inventory – Short Form Therapist version (WAI-T-C; Tracy and Kokotovic, 1989); Working Alliance Inventory Short Form Revised- Client version (WAI-SR; Hatcher and Gillaspy, 2006); Working Alliance Inventory Short Form Revised- Therapist version (WAI-SR-T; Hatcher and Gillaspy, 2006).

Outcome Measures: Beck Depression Inventory (BDI; Beck et al, 1961); Beck Depression Inventory – Second Edition (BDI-II; Beck & Brown, 1996); Center for the Epidemiological Studies of Depression-10 (CES-D-10; Andresen et al, 1994); Hamilton Rating Scale for Depression (HRSD; Hamilton, 1960); Montgomery-Asberg Depression Rating Scale (MADRS; Montgomery & Asberg, 1979).

Table 2: Effect sizes according to moderator

			Mean effect size				Heterogeneity			
Moderator	**Grouping**	***N***	**Mean *r***	**95% CI**	Z_r	p_r	***Q***	df_q	p_q	I^2
Time Point	Early	8	0.21	[.12 – .31]	4.22	0.000	1.80	7	0.97	0
	Mid	5	0.28	[.11 – .43]	3.25	0.001	1.64	4	0.800	0
	Late	6	0.27	[.13 – .40]	3.80	0.000	2.38	5	0.79	0
	Average	4	0.42	[.23 – .59]	4.01	0.000	10.2	3	0.02	70.5%
Rater	Client	9	0.27	[.19 – .36]	6.18	0.000	7.13	8	0.52	0
	Therapist	4	0.19	[.031 – .33]	2.35	0.02	1.58	3	0.66	0
	Observer	4	0.41	[.22 – .58]	3.95	0.000	6.62	3	0.085	54.7%