Behavioral activation treatments for depression in adults: a meta-analysis and review

Mazzucchelli T, Kane R, Rees C

CRD summary
The authors concluded that behavioural activation interventions were effective alternative treatments of depression in adults; further research was needed to determine the effects of different approaches. Potential for bias in the review and limitations with the included studies and analysis mean the authors’ conclusions should be interpreted with caution; interpretation should bear in mind that benefits were short term.

Authors’ objectives
To assess the effects of behavioural activation treatments on depression in adults.

Searching
PsycINFO and MEDLINE were searched between January 1970 and September 2008 for published and unpublished studies. No language restrictions were applied. Search terms were reported. Reference lists of retrieved articles were manually searched for additional articles.

Study selection
Randomised controlled trials (RCTs) that compared the effects of behavioural activation interventions with a control or other psychological or active pharmacological treatment on typically developing (without an intellectual disability) adults with depressive disorders or an elevated level of depressive symptomatology were eligible for inclusion. Behavioural activation interventions were categorised as pleasant activities, self control, contextual or behavioural activation treatment for depression (as defined in the review). Comparisons could include a control (waiting list or range of non-treatment options), cognitive behavioural therapy/cognitive therapy (CBT/CT) or other (such as psychotherapy, supportive counselling, assertiveness training, education, monitoring and increasing placebo activities, and usual treatment). The primary outcome of interest appeared to be change in depression, measured using instruments designed to explicitly measure symptoms of depression. Eligible studies were required to report sufficient data to extract effect sizes.

Most of the included studies were conducted in USA; others were conducted in Australia, Canada, The Netherlands and Spain. Patients were recruited from settings of hospital, university, community clinical or senior citizen apartment buildings. Mean age of patients ranged from 19 to 77 years (where reported). Where reported, intervention duration ranged from one to 20 weeks and the number of sessions ranged from one to 24 (each lasted between 20 and 120 minutes). Various scales and checklists were used to measure depression.

The authors did not state how many reviewers screened studies for inclusion.

Assessment of study quality
More than one reviewer assessed the quality of the included trials according to previously published nine-item criteria for psychotherapy research. Trials received a score between zero and 17. Disagreements were resolved through discussion. Trials were assessed for methodological rigour (including sample size, equivalence of comparison groups and treatment manual standards) according to the Task Force within Division 12 (Society of Clinical Psychology) of the American Psychological Society.

Data extraction
Means and standard deviations were extracted or estimated from relevant data (such as $X^2$ data or from t or F ratio) to calculate standardised mean difference effect sizes and their 95% confidence intervals (CIs). Where only diagnostic status data were reported, the arcsine transform method was used to adjust for dichotomisation. In trials that used more than one depression measure, the mean was calculated. In trials with two comparisons or where results were reported using subcategories (such as high/low depression severity), comparisons were combined taking into account relative proportions of participants in each condition.
The authors did not state how many reviewers performed the data extraction.

**Methods of synthesis**

Hedges' g effect sizes (adjusting for small sample bias) and 95% CIs were combined using a fixed-effect model where there was no statistical heterogeneity and a random-effects model where there was evidence of statistical heterogeneity. Statistical heterogeneity was assessed using Cochran's Q and $I^2$ statistics.

Linear regression or Pearson's correlation coefficient were used to investigate the influence of study characteristics on effect size. The relationship between interventions and measures of activity was investigated.

Subgroup analysis was undertaken for behavioural activation intervention approach (pleasant activities, self control, contextual or behavioural activation treatment for depression) and for patients with major depressive disorder. Sensitivity analyses were undertaken by removing trials in which control groups included a behavioural activation treatment.

Trials were assessed according to the Task Force within Division 12 (Society of Clinical Psychology) of the American Psychological Association's Division 12 Task Force. Follow-up ranged from post intervention up to 24 months.

**Results of the review**

Thirty-four RCTs (n=2,055) were included in the review. Attrition rates ranged from zero to 78%. Study quality was low in five RCTs (score of 5), moderate in 24 (score between 6 and 11) and high in five RCTs (score 12 to 15). Six RCTs met criteria required by the American Psychological Association's Division 12 Task Force. Follow-up ranged from post intervention up to 24 months.

**Behavioural activation versus control (16 RCTs, n=453):** At post treatment, behavioural activation interventions showed a significantly greater beneficial effect compared with controls using a random effects model (0.87, 95% CI 0.60 to 1.14, $I^2=43%$; 16 RCTs). Results remained significant for subgroup analysis for pleasant activities, self-control and contextual approaches; only one study assessed contextual approaches. There was some evidence of publication bias and the trim and fill method suggested that four studies were missing. The effects remained significant at one to three month follow-up with a fixed-effect model, but not when a random effects model was used ($I^2=77%$, five RCTs) and were not significant at seven to 12 months ($I^2=73%$, two RCTs).

**Behavioural activation versus CBT/CT (15 RCTs, n=536):** At post treatment there were no statistically significant differences between treatment groups and this remained the same in subgroup analysis. Sensitivity analysis did not alter the findings significantly. There was some evidence of publication bias, but the trim and fill method suggested no studies were missing. Results were similar between one and 24 months follow-up and remained non-significant.

**Behavioural activation versus psychotherapy or other interventions (17 RCTs, n=533):** At post treatment, behavioural activation interventions showed a medium effect size compared with other treatment (0.31, 95% CI 0.06 to 0.55, $I^2=58%$; 17 RCTs). Subgroup analysis showed similar findings for pleasant activities, but results were no longer significantly different for self-control and behavioural activation treatment for depression (although the authors reported that the effect sizes were medium to large). There was no evidence of selection bias. Follow-up between one and 12 months showed no statistically significant differences between treatment groups.

Linear regression or Pearson's correlation coefficient did not indicate that the study characteristics assessed influenced the findings (as reported in the review). Other results were reported in the review.

**Authors' conclusions**

Behavioural activation interventions may be considered a well-established and advantageous alternative to other treatments of depression in adults. However, further research is needed to determine whether simpler variations of behavioural activation are as effective as more complex variations.
CRD commentary

The review question and inclusion criteria were clearly defined. Outcomes of interest were not clearly stated. The literature search was somewhat limited, but included attempts to locate unpublished data and was not restricted by language. Formal assessment of publication bias indicated that bias was present for some comparisons. The authors acknowledged that a small number of studies individually assessed some of the behavioural activation interventions. The authors assessed some aspects of study quality, but this did not include assessment of criteria such as randomisation method, allocation concealment and blinding. The authors acknowledged the variable quality of the included studies. The authors did not state the process for study selection or data extraction, so reviewer error and bias could not be ruled out. There was evidence of statistical heterogeneity for some comparisons. Although the authors investigated statistical heterogeneity using appropriate methods, the causal factors were not identified; therefore, it was unclear whether pooling of the results was appropriate. Results were not reported for individual trials, which made it difficult to confirm the findings.

Given the potential for bias in the review and uncertainties about the included studies and analysis, the authors’ conclusions should be interpreted with caution. Any interpretation should bear in mind that the benefits were only short term.

Implications of the review for practice and research

Practice: The authors did not state any implications for practice.

Research: The authors stated that further long-term high-quality research was needed to investigate the different approaches used in behavioural activation interventions and the most effective components.

Funding

Not stated.

Bibliographic details


Original Paper URL


Indexing Status

Subject indexing assigned by CRD

MeSH

Adult; Behavior Therapy; Cognitive Therapy; Depressive Disorder, Major; Humans

AccessionNumber

12010001298

Date bibliographic record published

04/08/2010

Date abstract record published

09/02/2011

Record Status

This is a critical abstract of a systematic review that meets the criteria for inclusion on DARE. Each critical abstract contains a brief summary of the review methods, results and conclusions followed by a detailed critical assessment on the reliability of the review and the conclusions drawn.