Cost and Outcome of BehaviouRal Activation (COBRA): a randomised controlled trial of behavioural activation versus cognitive–behavioural therapy for depression

David A Richards,[1]* Shelley Rhodes,[1] David Ekers,[2] Dean McMillan,[3] Rod S Taylor,[1] Sarah Byford,[4] Barbara Barrett,[4] Katie Finning,[1] Poushali Ganguli,[4] Fiona Warren,[1] Paul Farrand,[5] Simon Gilbody,[3] Willem Kuyken,[6] Heather O'Mahen,[5] Ed Watkins,[5] Kim Wright,[5] Nigel Reed,[7] Emily Fletcher,[1] Steven D Hollon,[8] Lucy Moore,[1] Amy Backhouse,[1] Claire Farrow,[2] Julie Garry,[1] Deborah Kemp,[2] Faye Plummer,[9] Faith Warner[1] and Rebecca Woodhouse[3]

[1]University of Exeter Medical School, St Luke's Campus, Exeter, UK
[2]Psychological Therapy, Tees, Esk & Wear Valleys NHS Foundation Trust, County Durham, UK
[3]Department of Health Sciences, University of York, York, UK
[4]Institute of Psychiatry, Psychology & Neuroscience, Kings College London, London, UK
[5]Sir Henry Wellcome Building for Mood Disorders Research, University of Exeter, Exeter, UK
[6]Oxford Mindfulness Centre, Department of Psychiatry, University of Oxford, Oxford, UK
[7]Lived Experience Group, care of Sir Henry Wellcome Building for Mood Disorders Research, University of Exeter, Exeter, UK
[8]Department of Psychology, Vanderbilt University, Nashville, TN, USA
[9]Academic Unit of Elderly Care and Rehabilitation, Leeds Institute of Health Sciences, Bradford Royal Infirmary, Bradford, UK

*Corresponding author d.a.richards@exeter.ac.uk

Declared competing interests of authors: All authors report grants from the National Institute for Health Research (NIHR) during the course of the study. David A Richards reports grants from the European Science Foundation. David A Richards and Rod S Taylor have received funding support from NIHR Collaborations for Leadership in Applied Health Research and Care. David A Richards reports NIHR Clinical Development and Senior Clinical Fellowship and Senior Investigator Panel memberships. Rod S Taylor reports membership of NIHR Health Technology Assessment (HTA) programme themed call, NIHR HTA Efficient Study Designs Board and NIHR Health Services and Delivery Research Commissioning Boards. Simon Gilbody reports membership of the NIHR HTA Evidence Synthesis Board and NIHR HTA Efficient Study Designs Board. Willem Kuyken reports fees from Guilford Press for book royalties and Collaborative Case Conceptualisation.

Disclaimer: This report contains transcripts of interviews conducted in the course of the research and contains language that may offend some readers.

Published August 2017
DOI: 10.3310/hta21460

Scientific summary

Cost and Outcome of BehaviouRal Activation (COBRA)

Health Technology Assessment 2017; Vol. 21: No. 46
DOI: 10.3310/hta21460

NIHR Journals Library www.journalslibrary.nihr.ac.uk

Scientific summary

Background

Depression is a common, debilitating and costly disorder. Many patients request psychological therapy but the current best-evidenced therapy – cognitive–behavioural therapy (CBT) – is complex and costly. A simpler therapy, behavioural activation (BA), may be an effective alternative.

Objectives

1. To assess the clinical effectiveness of BA compared with CBT for depressed adults in terms of depression treatment response at 12 and 18 months.
2. To assess the cost-effectiveness of BA compared with CBT in terms of quality-adjusted life-years (QALYs) at 18 months.

We undertook a secondary process evaluation to investigate the moderating, mediating and procedural factors in BA and CBT that influence outcome.

Design

Randomised controlled non-inferiority trial.

Setting

Three English community mental health services.

Participants

Adults aged ≥ 18 years who met *Diagnostic and Statistical Manual of Mental Disorders*-Fourth Edition criteria for a major depressive disorder recruited from primary care and psychological therapy services in Devon, Durham and Leeds, excluding people who were receiving psychological therapy, were alcohol or drug dependent, were acutely suicidal or had attempted suicide in the previous 2 months, were cognitively impaired, had bipolar disorder, or who had psychosis or psychotic symptoms.

Randomisation

We randomly allocated participants in a 1 : 1 ratio to either BA or CBT arms stratified according to symptom severity on the Patient Health Questionnaire-9 (PHQ-9; < 19 vs. ≥ 19 points), antidepressant medication (ADM) use (yes/no) and recruitment site.

Allocation concealment

The registered Peninsula Clinical Trials Unit allocated participants remotely using a password-protected website after the researchers had collected and entered baseline data into a computer database.

Blinding

It was not possible to blind participants or clinicians. We ensured that research assessors were blind to participant allocation and we protected against assessment bias by using self-reported measures. We recorded instances where researchers were unblinded.

Interventions

After 5 days of training, NHS mental health workers (MHWs) and therapists delivered a maximum of 20 face-to-face weekly sessions of 1 hour duration of either BA or CBT, with the option of four additional booster sessions. MHWs and therapists received 1 hour of clinical supervision fortnightly from NHS psychological therapists clinically experienced in BA or CBT.

Behavioural activation

Behavioural activation, delivered by MHWs at NHS Agenda for Change (AfC) band 5 grade, was a structured programme increasing contact with potentially antidepressant environmental reinforcers and reducing the frequency of negatively reinforced avoidant behaviours. Specific BA techniques included the use of a functional analytical approach, self-monitoring, identifying 'depressed behaviours', developing alternative goal-orientated behaviours and scheduling. The role of avoidance and rumination was addressed through functional analysis and alternative response development.

Cognitive–behavioural therapy

Cognitive–behavioural therapy, delivered by NHS AfC band 7 therapists, was a structured programme to identify and modify negative automatic thoughts, maladaptive beliefs and, if indicated, underlying core beliefs. Specific CBT techniques included scheduling activity and mastery behaviours, the use of thought records and modifying maladaptive beliefs and rumination content. The behavioural elements in CBT focused on increasing activity with behavioural experiments to test specific cognitive beliefs rather than the contextual, functional analytical approach of the BA trial arm.

Measures

Baseline information

We collected demographic data at baseline on gender, age, ethnic origin, education level, employment, marital status, number of children, presence and duration of ADM treatment, previous history and age at onset of depression, and presence of comorbid anxiety disorder(s).

Primary clinical outcome

Depression severity (as measured via the PHQ-9) at 12 months.

Secondary clinical outcome

Major depressive disorder status; number of depression-free days; anxiety (as measured via the Generalised Anxiety Disorder-7 questionnaire); health-related quality of life (as measured via the Short Form questionnaire-36 items) at 6, 12 and 18 months; PHQ-9 at 6 and 18 months.

Economic outcomes

Cost per QALY at 18 months post randomisation, derived from the EuroQol-5 Dimensions, three-level version. We collected resource use data associated with delivery of BA and CBT from clinical records.

We measured all other health and social care services used, including medication prescription using the adult service use schedule. We measured productivity losses using the absenteeism and presenteeism questions of the World Health Organization's Health and Work Performance Questionnaire.

Process data

Behaviour (Behavioural Activation for Depression Scale); beliefs (Dysfunctional Attitudes Scale); rumination (Ruminative Response Scale); hedonic tone (Snaith–Hamilton Pleasure Scale); per protocol (PP) treatment adherence (from therapist case records); qualitative data via semistructured interviews to assess acceptability of BA and CBT for participants and clinicians.

Adverse events

Deaths from whatever cause and all self-harm and suicide attempts.

Sample size

We powered the trial at 90% ($\alpha = 0.05$) to detect a non-inferiority margin of 1.9 PHQ-9 points, inflating our sample size by 20% for participant attrition. Consequently, we needed to recruit 440 participants, 220 per arm, to detect a between-group non-inferiority margin of 1.90 in PHQ-9 points at one-sided 2.5% alpha.

Statistical methods and analyses

Clinical outcomes

We assessed equivalence of baseline characteristics and outcomes in the two groups descriptively. We analysed primary and secondary outcomes in accordance with Consolidated Standards of Reporting Trials guidelines for non-inferiority and equivalence trials, undertaking both intention-to-treat (ITT) and PP analyses. We compared observed primary and secondary outcomes between groups 12 months after randomisation using linear regression models adjusted for baseline outcome values and stratification/minimisation variables. We extended primary analysis models to fit interaction terms to explore differences in treatment effect from baseline symptom severity and ADM usage. We undertook secondary analyses to compare groups at follow-up across 6, 12 and 18 months using mixed-effects repeated measures regression. We ran sensitivity analyses for both primary and secondary analyses to assess the impact of missing data using multiple imputation models. We calculated the relative proportions of participants meeting criteria for 'recovery' (proportions of participants with PHQ-9 scores of ≤ 9 points) and 'response' (50% reduction in PHQ-9 scores from baseline).

Economic outcomes

We took the UK NHS and Personal Social Services perspective consistent with the UK National Institute for Health and Care Excellence (NICE)'s reference case and examined a broader societal perspective, adding productivity losses attributable to time off work, in a sensitivity analysis. We compared the costs and cost-effectiveness of BA and CBT at the final 18-month follow-up to capture the impact of events such as relapse, with unit costs from the 2013–14 financial year. We assessed cost-effectiveness in terms of QALYs using the net benefit approach. We analysed differences in mean cost per participant at 18 months using parametric *t*-tests, with the validity of results confirmed using bias-corrected, non-parametric bootstrapping. We calculated incremental cost-effectiveness ratios (ICERs) and constructed cost-effectiveness planes using 1000 bootstrapped resamples from regression models of total health- and social-care costs and outcome by treatment group, using these replications to calculate the probability that each treatment is the optimal choice for different values a decision-maker is willing to pay for 1-unit outcome of improvement. We

produced cost-effectiveness acceptability curves (CEACs) illustrating the probability that BA is cost-effective compared with CBT, which is dependent on willingness to pay per QALY. We controlled for stratification variables and baseline values of the variables of interest, truncating data to exclude influential outliers.

Process outcomes

Interactions between treatment allocation and each process covariate were investigated at 6, 12 and 18 months' follow-up for PHQ-9. A series of models were performed, adjusting for the stratification variables, trial site, baseline ADM use and baseline PHQ-9 score. Each model included the specific covariate being investigated as a potential moderator and its interaction with treatment allocation. For mediation, we used a structural equation modelling approach to evaluate the effect of each individual mediator at each follow-up time, on the primary outcome. We included all mediators measured at a specific follow-up time in an overall model for each follow-up point. Qualitative data were analysed using a framework analysis combining deductive themes from the topic guides and inductive themes emerging from the data. Transcripts were examined thematically across the whole data set as well as in the context of each interview, using constant comparative techniques.

Results

We recruited 440 participants, randomly allocating 221 (50%) to the BA group and 219 (50%) to the CBT group. Patient- and trial-level characteristics at baseline were well balanced between groups. Participants received a mean of 11.5 [standard deviation (SD) 7.8] BA sessions or 12.5 (SD 7.8) CBT sessions. We found that BA was non-inferior to CBT [ITT: CBT 8.4 PHQ-9 points (SD 7.5 PHQ-9 points), BA 8.4 PHQ-9 points (SD 7.0 PHQ-9 points), mean difference 0.1, 95% confidence interval (CI) –1.3 to 1.5 PHQ-9 points, $p = 0.89$; PP: CBT 7.9 PHQ-9 points (SD 7.3 PHQ-9 points); BA 7.8 PHQ-9 points (SD 6.5 PHQ-9 points), mean difference 0.0, 95% CI –1.5 to 1.6 PHQ-9 points, $p = 0.99$]. Between 61% and 70% of ITT and PP participants in both groups met criteria for recovery or response, with no difference in the proportions of patients in each group. We found no difference between groups on secondary outcomes at any time point. All findings were robust to sensitivity analyses.

Two (1%) non-trial-related deaths [one (1%) multidrug toxicity in the BA group and one (1%) cancer in the CBT group] and 15 depression-related, but not treatment-related, serious adverse events (three in the BA group and 12 in the CBT group) occurred in three (2%) participants in the BA group [two (1%) patients who overdosed and one (1%) who self-harmed] and eight (4%) participants in the CBT group [seven (4%) who overdosed and one (1%) who self-harmed].

We found a significant difference in mean intervention costs between the two groups, but no differences in other categories of cost or in total health- and social-care costs. As costs were lower and QALY outcomes better in the BA group than in the CBT group, this generated an ICER of –£6865, suggesting that BA dominates CBT (i.e. is both cheaper and more effective). The CEAC showing the probability of BA being cost-effective compared with CBT does not fall below 75% and is closer to 80% at standard NICE-preferred willingness-to-pay levels of £20,000–30,000 per QALY. All findings were robust to sensitivity analyses.

We found a weak moderating effect of baseline PHQ-9 score on treatment effect, with regard to PHQ-9 at 12 and 18 months' follow-up, indicating that BA may be a better choice of treatment for patients with higher baseline PHQ-9 scores. The only significant mediation effects were that overall treatment fidelity mediated the effect of treatment on PHQ-9 at 12 months' follow-up, with basic and overall treatment fidelity mediating the effect of treatment on PHQ-9 at 18 months' follow-up. Qualitative data showed that, despite being challenging at times, BA and CBT were acceptable and feasible for participants, MHWs and therapists, and effected changes in people's specific symptoms and in their lives more broadly. Despite experiencing initial difficulties that could be detected by some participants, with sufficient training, experience and supervision, junior MHWs could feel confident in delivering BA effectively.

Conclusions

Behavioural activation for depression is not inferior to CBT in terms of reduction of depression symptoms and is cost-effective compared with CBT against commonly applied decision-maker willingness-to-pay thresholds. We observed our results using both ITT and PP analyses, using a conservative non-inferiority margin. Our results in both groups compare favourably with a meta-analysis of the effects of CBT that estimate proportions of patients with remissions of around 50%. Our economic outcomes were driven by the lower costs of the MHWs who delivered BA, compared with the more experienced psychological therapists who routinely deliver CBT. Our study results, therefore, substantiate the hypothesis that BA is as effective as CBT and that BA's simplicity renders it suitable for delivery by junior MHWs with no professional training in psychological therapies.

Baseline PHQ-9 score had a weak moderating effect on depression symptoms at 12 and 18 months' follow-up, this interaction effect indicating that BA may be a better choice of treatment for patients with higher baseline PHQ-9 scores. We found that only treatment fidelity reliably showed an interaction with outcome, demonstrating the importance of MHWs and therapists adhering to clinical protocols. BA and CBT were both acceptable and feasible for participants, MHWs and therapists. Importantly, junior MHWs can deliver BA effectively, although they need training, experience and supervision to feel confident in delivering BA.

Strengths and limitations

To date, COBRA is the largest trial of BA and one of the largest psychological treatment trials for depression. We followed up participants for 18 months and our economic analysis is one of few in this field. Therapists and MHWs working in three different routine NHS settings delivered treatment, providing evidence of potential generalisability. We could not mask patients or clinicians to treatment allocation, but used self-reported outcome measures and robust researcher-masking procedures to reduce unmasking to $< 5\%$.

In this pragmatic trial many depressed participants in both groups were also taking ADM, although most had been doing so for a considerable time before entering the trial. Our levels of attrition and outcome loss to follow-up were low, similar to other trials in this area, but are still a limitation. However, our between-group inferences were robust to data imputation. Around one-third of participants chose not to complete a PP dose of treatment, a finding common in both psychotherapy trials and routine practice.

Implications

For years, CBT has been the foremost psychological therapy recommended by therapists, researchers and policy-makers. Our results challenge this dominance and suggest that BA could be a front-line treatment for depression. Our most striking finding is that BA leads to comparable clinical outcomes for patients with depression, but at a financial saving to clinical providers of 21% compared with the cost of provision of CBT, with no compensatory use of other health-care services by patients. There are substantial implications for the scalability of psychological treatment for depression in the UK and internationally, given the greater availability and ease with which a BA workforce could be trained than could a CBT workforce.

Although many obstacles exist to successful dissemination in addition to training of MHWs, our findings suggest that health services globally could reduce the need for costly professional training and infrastructure, reduce waiting times and increase access to psychological therapies. Our findings have substantial implications given the increasing global pressure for cost-containment across health systems in high-income countries, and the need to develop accessible, scalable interventions in low- and middle-income countries.

Our results, therefore, offer hope to many societies, cultures and communities worldwide, rich and poor, struggling with the effect of depression on the health of their people and economies.

Trial registration

This trial is registered as ISRCTN27473954.

Funding

Funding for this study was provided by the Health Technology Assessment programme of the National Institute for Health Research.

Health Technology Assessment

HTA/HTA TAR

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Impact factor: 4.236

Health Technology Assessment is indexed in MEDLINE, CINAHL, EMBASE, The Cochrane Library and the Clarivate Analytics Science Citation Index.

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: journals.library@nihr.ac.uk

The full HTA archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hta. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Health Technology Assessment* journal

Reports are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

HTA programme

The HTA programme, part of the National Institute for Health Research (NIHR), was set up in 1993. It produces high-quality research information on the effectiveness, costs and broader impact of health technologies for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined as all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

For more information about the HTA programme please visit the website: http://www.nets.nihr.ac.uk/programmes/hta

This report

The research reported in this issue of the journal was funded by the HTA programme as project number 10/50/14. The contractual start date was in April 2012. The draft report began editorial review in October 2016 and was accepted for publication in March 2017. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).