This is a repository copy of Cost and Outcome of Behavioural Activation versus Cognitive Behavioural Therapy for Depression (COBRA) : a randomised, controlled, non-inferiority trial.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/105790/

Version: Published Version

Article:
Richards, David A, Ekers, David, McMillan, Dean et al. (15 more authors) (2016) Cost and Outcome of Behavioural Activation versus Cognitive Behavioural Therapy for Depression (COBRA) : a randomised, controlled, non-inferiority trial. The Lancet. pp. 871-880. ISSN 0140-6736

https://doi.org/10.1016/S0140-6736(16)31140-0

Reuse
This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:
https://creativecommons.org/licenses/

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.
Cost and Outcome of Behavioural Activation versus Cognitive Behavioural Therapy for Depression (COBRA): a randomised, controlled, non-inferiority trial

David A Richards, David Ekers, Dean McMillan, Rod S Taylor, Sarah Byford, Fiona C Warren, Barbara Barrett, Paul A Farrand, Simon Gilbody, Willem Kuyken, Heather O’Mahen, Ed R Watkins, Kim A Wright, Steven D Hollon, Nigel Reed, Shelley Rhodes, Emily Fletcher, Katie Finning

Summary

Background Depression is a common, debilitating, and costly disorder. Many patients request psychological therapy, but the best-evidenced therapy—cognitive behavioural therapy (CBT)—is complex and costly. A simpler therapy—behavioural activation (BA)—might be as effective and cheaper than is CBT. We aimed to establish the clinical efficacy and cost-effectiveness of BA compared with CBT for adults with depression.

Methods In this randomised, controlled, non-inferiority trial, we recruited adults aged 18 years or older meeting Diagnostic and Statistical Manual of Mental Disorders IV criteria for major depressive disorder from primary care and psychological therapy services in Devon, Durham, and Leeds (UK). We excluded people who were receiving psychological therapy, were alcohol or drug dependent, were acutely suicidal or had attempted suicide in the previous 2 months, or were cognitively impaired, or who had bipolar disorder or psychosis or psychotic symptoms. We randomly assigned participants (1:1) remotely using computer-generated allocation (minimisation used; stratified by depression severity [Patient Health Questionnaire 9 (PHQ-9) score of ≤19 vs ≥19], antidepressant use, and recruitment site) to BA from junior mental health workers or CBT from psychological therapists. Randomisation was done at the Peninsula Clinical Trials Unit was concealed from investigators. Treatment was assigned open label, but outcome assessors were masked. The primary outcome was depression symptoms according to the PHQ-9 at 12 months. We analysed all those who were randomly allocated and had complete data (modified intention to treat [mITT]) and also all those who were randomly allocated, had complete data, and received at least eight treatment sessions (per protocol [PP]). We analysed safety in the mITT population. The non-inferiority margin was 1·9 PHQ-9 points. This trial is registered with the ISRCTN registry, number ISRCTN27473954.

Findings Between Sept 26, 2012, and April 3, 2014, we randomly allocated 221 (50%) participants to BA and 219 (50%) to CBT. 175 (79%) participants were assessable for the primary outcome in the mITT population in the BA group compared with 189 (86%) in the CBT group, whereas 135 (61%) were assessable in the PP population in the BA group compared with 151 (69%) in the CBT group. BA was non-inferior to CBT (mITT: CBT 8·4 PHQ-9 points [SD 7·5]; BA 8·4 PHQ-9 points [7·0], mean difference 0·1 PHQ-9 points [95% CI –1·3 to 1·5], p=0·89; PP: CBT 7·9 PHQ-9 points [6·5]; BA 7·8 PHQ-9 points [6·5], mean difference 0·1 PHQ-9 points [–1·9 to 1·9], p=0·89). 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 non-trial-related, serious adverse events (three [1.3%] in the BA group and 12 [5.5%] 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).

Interpretation We found that BA, a simpler psychological treatment than CBT, can be delivered by junior mental health workers with less intensive and costly training, with no lesser effect than CBT. Effective psychological therapy for depression can be delivered without the need for costly and highly trained professionals.

Funding National Institute for Health Research.

Copyright © The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY license.

Introduction Clinical depression is a common and debilitating mental health disorder, being the second largest cause of global disability. Globally, the effect of depression on aggregate economic output is predicted to be US$5·36 trillion between 2011 and 2030. Reduction of these substantial costs is a key objective for low-income, middle-income, and high-income countries alike. Antidepressant medication and cognitive behavioural therapy (CBT) have the most clinical evidence. However, although antidepressant medications are cheap, their use is limited by side-effects, poor patient adherence, and discontinuation relapse risk. CBT is as effective as are antidepressants and provides long-term protection against relapse, but it is complex and its effectiveness is dependent on the skills of...
Research in context

Evidence before this study
Authors of published systematic reviews, including a Cochrane review, have commented on the limitations of existing evidence for the effectiveness of behavioural activation (BA) for depression compared with cognitive behavioural therapy (CBT) and the scarcity of cost-effectiveness data, with the existing evidence insufficiently robust to establish comparability. Authors of the Cochrane review called for studies that improve the quality of evidence. Our pretrial evidence took published review findings from the UK National Institute for Health and Care Excellence (NICE), who reported no difference in treatment outcome between BA and CBT immediately after treatment (Hedges’ g 0·129 [95% CI −0·400 to 0·122]; p=0·296) and subsequent follow-up (0·135 [−0·456 to 0·185]; p=0·409). The authors of NICE’s review regarded the existing international evidence as insufficient to recommend BA for first-line treatment in clinical guidelines for depression.

Added value of this study
This trial addresses these research recommendations and is, to our knowledge, the only high-quality, fully powered non-inferiority and cost-effectiveness study addressing both the effects and costs of BA compared with CBT for depression. When we combine the data from our study with data from other international studies in the meta-analysis done by NICE, our data reduce the 95% CIs around the effect size for depression symptoms immediately after treatment (Hedges’ g 0·054 [95% CI −0·214 to 0·107]; p=0·514) and at follow-up (0·059 [−0·234 to 0·115]; p=0·503) and unequivocally show both non-inferiority of BA compared with CBT and that BA is more cost-effective than is CBT against commonly applied decision maker willingness-to-pay thresholds.

Implications of all the available evidence
Junior mental health workers with no professional training in psychological therapies can deliver behavioural activation, a simple psychological treatment, with no lesser effect than CBT has and at less cost. Effective psychological therapy for depression can be delivered without the need for costly and highly trained professionals.

Methods
Study design and participants
In this randomised, controlled, open-label, non-inferiority trial (the Cost and Outcome of Behavioural Activation versus Cognitive Behaviour Therapy for Depression [COBRA] trial), we recruited participants from primary care and psychological therapy services in Devon, Durham, and Leeds (UK). Eligible participants were adults aged 18 years or older who met diagnostic criteria for major depressive disorder assessed by researchers using a standard clinical interview (Structured Clinical Interview for the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition [SCID]®). We excluded people at interview who were receiving psychological therapy, were alcohol or drug dependent, were acutely suicidal or had attempted suicide in the previous 2 months, or were cognitively impaired, or who had bipolar disorder or psychosis or psychotic symptoms.

We recruited participants by searching the electronic case records of general practices and psychological therapy services for patients with depression, identifying potential participants from depression classification codes. Practices or services contacted patients to seek permission for researcher contact. The research team interviewed those that responded, provided detailed information on the study, took informed written consent, and assessed people for eligibility. The UK South West Research Ethics Committee gave national approval for the study (NRES/07/H1208/60). The protocol has been published previously.

Randomisation and masking
After eligibility was established, consent agreed, and baseline data collected, we randomly allocated participants...
(1:1) to BA or CBT using computer-generated allocation, stratified by depression severity according to the Patient Health Questionnaire 9 (PHQ-9)\(^a\) (<19 vs ≥19), antidepressant use (taking antidepressives or not), and recruitment site (Devon, Durham, or Leeds). A computer-based system allocated the first 20 participants to each group on a truly random basis. For subsequent participants, allocation was minimised to maximise the likelihood of balance in stratification variables across the two study groups. The registered Peninsula Clinical Trials Unit (Plymouth University, Plymouth, UK) allocated participants remotely after baseline data entry to ensure allocation concealment. Treatment was given open label, but outcome assessors were masked to participants' allocations. Concealment was ensured by use of an externally administered password-protected trial website with retention of a stochastic element to the minimisation algorithm. We recorded instances when outcome assessors were unmasked during interviews if participants informed them of their allocation.

Procedures

We developed our clinical protocols in line with published treatment protocols\(^b,c\) including those from our own trials\(^d,e\), advice from international collaborators, and NICE recommendations\(^f\) for duration and frequency of BA and CBT. Full-time National Health Service (NHS) MHWs and therapists worked half of their working week for COBRA (with the other half worked as normal) and followed written manuals to deliver a maximum of 20 sessions over 16 weeks, with the option of four additional booster sessions if the patients wanted them.\(^g\)

Treatment included core and supplementary techniques appropriate to the BA or CBT protocol to be used as clinically indicated; for example, behavioural or cognitive strategies for management of anxiety. All core components of both treatments were delivered by session eight, which we considered to represent a minimally sufficient dose of therapy (appendix). Sessions were face to face, lasting for 60 min. BA and CBT experts on the trial team trained MHWs and therapists for 5 days in face-to-face, lasting for 60 min. BA and CBT experts on the trial team trained MHWs and therapists for 5 days in trial team trained MHWs and therapists for 5 days in either BA or CBT. MHWs and therapists were assessed competently at the end of training with use of standardised quality criteria instruments consistent with for competence at the end of training with use of a computer-generated random number sequence) sample of tapes, stratified by therapist, therapy session, and intervention, for competence using the Revised Cognitive Therapy Scale\(^h\) for CBT (range 0–72) and the Quality of Behavioral Activation Scale (Dimidjian S, University of Colorado, personal communication) for BA (range 0–96). All therapists recorded the specific therapeutic techniques that they had used for each session on a checklist.

Outcomes

The primary outcome was self-reported depression severity (PHQ-9 score\(^i\)) at 12 months. Secondary outcomes were PHQ-9 score at 6 months and 18 months and Diagnostic and Statistical Manual of Mental Disorders IV major depressive and anxiety disorder status and number of depression-free days between follow-ups (SCID),\(^j\) anxiety (Generalized Anxiety Disorder 7),\(^k\) and health-related quality of life (36-Item Short Form Survey)\(^l\) at 6 months, 12 months, and 18 months. In adverse events, we recorded deaths from whatever cause and all self-harm and suicide attempts. The independent Data Management Committee reviewed all adverse events and made relevant trial conduct recommendations.

Statistical analysis

Previous research has suggested that non-inferiority margins should be half of the mean controlled effect size.
from historical trials. Accordingly, we estimated the non-inferiority margin for the primary outcome using meta-analysis data from trials of BA for which BA was superior to controls by a mean of 0·7 SD units (95% CI 0·39–1) or 3·8 PHQ-9 score units (2·1–5·4). Therefore, our non-inferiority margin was 1·9 PHQ-9 points (ie, 0·5 × 3·8). We inflated our sample size by 20% for participant follow-up attrition. We planned to recruit

Figure 1: Trial profile
(A) 6 month, (B) 12 month, and (C) 18 month follow-up. BA=behavioural activation. CBT=cognitive behavioural therapy. *Includes four participants who were initially allocated in error and subsequently excluded.
220 participants per arm to detect a between-group non-inferiority margin of 1.9 PHQ-9 points with a one-sided 2.5% α. Furthermore, although findings from a trial of CBT have shown little effect of outcome clustering by therapists, the presence of a small therapist clustering effect (ie, an intraclass correlation coefficient of 0-01) would still provide the same power.

We did all analyses using a statistical analysis plan prepared in the first 6 months of the trial, agreed with the Trial Management Group, Trial Steering Committee, and Data Management Committee. We assessed between-group equivalence of baseline characteristics and outcomes descriptively and did a descriptive analysis of baseline characteristics by recruitment method.

We compared observed primary and secondary outcomes between groups 12 months after randomisation using linear regression models adjusted for baseline outcome values and stratification variables. We did multiple imputation models. We did sensitivity analyses for different definitions of PP (eight, 12, 16, and 20 treatment sessions) to check security of our primary outcome and for different definitions of PP in the mITT population. We did sensitivity analyses for our primary outcome and for different definitions of PP (eight, 12, 16, and 20 treatment sessions) to check security of inference of non-inferiority.

We accepted non-inferiority of BA to CBT (in a 0-0.025 level test) if the lower bound of the two-sided 95% CI (equivalent to the upper bound of one-sided 97-5% CI) was within the non-inferiority margin of –1.9 PHQ-9 points. We checked for non-equivalence of the primary outcome at all follow-up points using the same approach.

We did secondary analyses to compare groups at follow-up across 6 months, 12 months, and 18 months using hierarchical linear regression. To ease clinical interpretation, we calculated proportions of recovery (participants with PHQ-9 scores of ≤9) and response (50% reduction from baseline PHQ-9 scores). We ran sensitivity analyses to assess the likely effect of missing data using multiple imputation models. We did imputation by treatment group using chained equations to create 20 complete datasets under the assumption that data were missing at random. Imputation models included covariates as defined for the primary analysis model and auxiliary variables that were predictive of outcomes. After analysis, we combined the effect estimates from the imputed datasets using Rubin’s rule. For economic analyses, we took the UK NHS and personal social services perspective consistent with the NICE reference case, also examining a wide societal perspective, adding productivity losses due to time off work in a sensitivity analysis. We collected participants’ use of BA and CBT from clinical records, with additional resource information (eg, training, supervision, and other non-face-to-face activities) from therapists and trainers. We used the Adult Service Use Schedule to measure other health and social care services used, including psychotropic medications. We measured productivity losses using the absenteeism and presenteeism questions from the Adult Service Use Schedule to measure other health and social care services used, including psychotropic medications. We measured productivity losses using the absenteeism and presenteeism questions from the

<table>
<thead>
<tr>
<th>Trial characteristics</th>
<th>BA (n=221)</th>
<th>CBT (n=239)</th>
<th>All (n=440)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient characteristics</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td>43.9 (14.1)</td>
<td>43.0 (14.1)</td>
<td>43.5 (14.1)</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>79 (26%)</td>
<td>71 (32%)</td>
<td>150 (34%)</td>
</tr>
<tr>
<td>Female</td>
<td>142 (64%)</td>
<td>148 (68%)</td>
<td>290 (66%)</td>
</tr>
<tr>
<td>Number of episodes of depression (including current)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>7.0 (15.0)</td>
<td>6.3 (13.8)</td>
<td>6.7 (14.4)</td>
</tr>
<tr>
<td>Median</td>
<td>3.0 (1-5)</td>
<td>2.0 (1-5)</td>
<td>3.0 (1-5)</td>
</tr>
<tr>
<td>Age of onset of first depression episode (years)</td>
<td>27.2 (15.0)</td>
<td>26.3 (13.5)</td>
<td>26.7 (14.2)</td>
</tr>
<tr>
<td>Duration of antidepressant treatment (weeks)^a</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean, n</td>
<td>215 (81%)</td>
<td>160 (64%)</td>
<td>164 (66%)</td>
</tr>
<tr>
<td>Median, n</td>
<td>210 (78)</td>
<td>168 (72)</td>
<td>198 (76)</td>
</tr>
<tr>
<td>At least one comorbid anxiety disorder</td>
<td>131 (59%)</td>
<td>141 (64%)</td>
<td>272 (62%)</td>
</tr>
<tr>
<td>Marital status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>68 (31%)</td>
<td>59 (27%)</td>
<td>127 (29%)</td>
</tr>
<tr>
<td>Cohabiting (not married)</td>
<td>29 (13%)</td>
<td>25 (11%)</td>
<td>54 (12%)</td>
</tr>
<tr>
<td>Civil partnership</td>
<td>1 (&lt;1%)</td>
<td>1 (&lt;1%)</td>
<td>2 (&lt;1%)</td>
</tr>
<tr>
<td>Married</td>
<td>84 (38%)</td>
<td>92 (42%)</td>
<td>176 (40%)</td>
</tr>
<tr>
<td>Divorced or separated</td>
<td>39 (18%)</td>
<td>42 (19%)</td>
<td>81 (18%)</td>
</tr>
<tr>
<td>Number of children</td>
<td>0</td>
<td>74 (33%)</td>
<td>146 (33%)</td>
</tr>
<tr>
<td>1</td>
<td>35 (16%)</td>
<td>31 (14%)</td>
<td>66 (15%)</td>
</tr>
<tr>
<td>2</td>
<td>67 (30%)</td>
<td>69 (32%)</td>
<td>136 (31%)</td>
</tr>
<tr>
<td>3</td>
<td>31 (14%)</td>
<td>27 (12%)</td>
<td>58 (13%)</td>
</tr>
<tr>
<td>&gt;4</td>
<td>14 (6%)</td>
<td>20 (9%)</td>
<td>34 (8%)</td>
</tr>
<tr>
<td>Level of education</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No qualifications</td>
<td>25 (11%)</td>
<td>39 (14%)</td>
<td>55 (12%)</td>
</tr>
<tr>
<td>GCSEs or O Levels</td>
<td>36 (16%)</td>
<td>43 (20%)</td>
<td>79 (18%)</td>
</tr>
<tr>
<td>AS or A Levels</td>
<td>28 (13%)</td>
<td>22 (10%)</td>
<td>50 (11%)</td>
</tr>
<tr>
<td>NVQ or other vocational qualification</td>
<td>54 (24%)</td>
<td>71 (32%)</td>
<td>125 (28%)</td>
</tr>
<tr>
<td>Undergraduate degree</td>
<td>44 (20%)</td>
<td>35 (16%)</td>
<td>79 (18%)</td>
</tr>
<tr>
<td>Postgraduate degree</td>
<td>28 (13%)</td>
<td>14 (6%)</td>
<td>42 (10%)</td>
</tr>
<tr>
<td>Doctoral degree</td>
<td>2 (1%)</td>
<td>1 (&lt;1%)</td>
<td>3 (1%)</td>
</tr>
<tr>
<td>Professional degree (eg, MD)</td>
<td>4 (2%)</td>
<td>3 (1%)</td>
<td>7 (2%)</td>
</tr>
</tbody>
</table>

| Ethnicity             |           |             |             |
| White British         | 204 (92%) | 197 (90%)   | 401 (91%)   |
| Other                 | 17 (8%)   | 22 (10%)    | 39 (9%)     |

(Number of episodes of depression (including current) | Mean, n | Median, n | At least one comorbid anxiety disorder | Marital status | Number of children | Level of education | Ethnicity | Articles | www.thelancet.com Vol 388 August 27, 2016 875 | (Table 1 continues on next page) |
Table 1: Baseline characteristics

<table>
<thead>
<tr>
<th>PHQ-9 category</th>
<th>BA (n=221)</th>
<th>CBT (n=219)</th>
<th>All (n=440)</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;19</td>
<td>118 (53%)</td>
<td>118 (54%)</td>
<td>236 (54%)</td>
</tr>
<tr>
<td>≥19</td>
<td>103 (47%)</td>
<td>101 (46%)</td>
<td>204 (46%)</td>
</tr>
<tr>
<td>Antidepressant use</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>172 (78%)</td>
<td>173 (79%)</td>
<td>345 (78%)</td>
</tr>
<tr>
<td>No</td>
<td>49 (22%)</td>
<td>46 (21%)</td>
<td>95 (22%)</td>
</tr>
<tr>
<td>Site</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Devon</td>
<td>74 (33%)</td>
<td>73 (33%)</td>
<td>147 (33%)</td>
</tr>
<tr>
<td>Durham</td>
<td>79 (36%)</td>
<td>78 (36%)</td>
<td>157 (36%)</td>
</tr>
<tr>
<td>Leeds</td>
<td>68 (31%)</td>
<td>68 (31%)</td>
<td>136 (31%)</td>
</tr>
</tbody>
</table>

Data are n (%), mean (SD), or median (IQR), unless otherwise indicated. IAPT=Improving Access to Psychological Therapies. GCSE=General Certificate of Secondary Education. O Level=Ordinary Level. AS Level=Advanced Subsidiary Level. A Level=Advanced Level. NVQ=National Vocational Qualification. MD=Doctor of Medicine. PHQ-9=Patient Health Questionnaire. *16 participants who reported that they were using antidepressant medication at baseline did not report duration of use (12 in the BA group and four in the CBT group).

Results

Between Sept 26, 2012, and April 3, 2014, we recruited 440 participants, randomly allocating 221 (50%) to the BA group and 219 (50%) to the CBT group (figure 1). 175 (79%) participants were assessable for the primary outcome in the mITT population in the BA group compared with 189 (86%) in the CBT group, whereas 135 (61%) were assessable in the PP population in the BA group compared with 151 (69%) in the CBT group. We noted no evidence of a difference in patient characteristics between recruitment methods (appendix). Patient-level and trial-level characteristics at baseline were well balanced between groups (table 1). PHQ-9 score at baseline was negatively skewed, with a high proportion of participants scoring towards the upper end of the distribution (data not shown), but scores were similar between groups (table 2).

Ten MHWs provided BA (median 22 participants each (IQR 19–25)) and 12 therapists provided CBT (21 [13–23]). MHWs had a mean of 18 months mental health experience (SD 11) and CBT therapists had a mean of 22 months post-CBT qualification (24). We removed one CBT therapist from the trial in the early stages who did not meet acceptable competency. Participants received a mean of 11.5 BA sessions (7–8) or 12.5 CBT sessions (7–8). 305 participants (69%) completed the PP number of at least eight sessions (BA 147 [67%] patients, mean 16.1 sessions [SD 5.3]; CBT 158 [72%] patients, mean 16.4 sessions [5.4]); participants completing less than eight sessions (135 [31%]; BA 74 [33%], CBT 61 [28%]) completed a mean of 2.5 BA sessions (SD 1.9) or 2.6 CBT sessions (2.1). MHWs and therapists met acceptable competency standards: mean Quality of Behavioral Health and Work Performance Questionnaire.25 We calculated effectiveness in terms of quality-adjusted life-years (QALYs) using the EuroQol-5D-3L measure of health-related quality of life.26 We assigned health states from the EuroQol-5D-3L measure a utility score using responses from a representative sample of adults in the UK.27 We calculated QALYs as the area under the curve defined by the utility values at baseline and each follow-up, assuming that utility score changes over time followed a linear path.

We compared the costs and cost-effectiveness of treatments at 18 months to capture the economic effect of events like relapse with unit costs from the 2013–14 financial year.28,29 We discounted costs and QALYs in year 2 at 3.5%.30 We used complete case analysis with missing data explored in a sensitivity analysis using multiple imputation with chained equations. We calculated the cost of each treatment using a microcosting (bottom-up) approach.31 We based MHW costs on NHS Agenda for Change salary band five (salary range £21909–28462; US$31662–41130; €27726–35993) for BA and band seven (£31383–41373; US$45350–59786; €39738–52388) for CBT therapists and included employer National Insurance and pension contributions plus capital, administrative, and managerial costs. We calculated cost per h using standard working time assumptions,32 weighted to account for time spent on non-patient-facing activities. We applied nationally applicable unit costs for other health and social care services.

We assessed cost-effectiveness in terms of QALYs using the net benefit approach.33 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.34 We calculated incremental cost-effectiveness ratios and constructed cost-effectiveness planes using 1000 bootstrapped resamples from regression models of total cost and outcome by treatment group. We used these bootstrapped replications to calculate the probability that each of the treatments is the optimal choice for different values a decision maker is willing to pay for a unit improvement in outcome, representing uncertainty around the cost and effectiveness estimates, with cost-effectiveness acceptability curves illustrating the probability that BA is cost-effective compared with CBT, dependent on willingness to pay per QALY.35 We controlled for stratification variables and baseline values of the variables of interest, truncating data to exclude influential outliers—ie, cases with total costs in the 99th percentile that make a significant difference to the results. We did all analyses using Stata version 14.1. This trial is registered with the ISRCTN registry, number ISRCTN27473954.

Role of the funding source

The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. DAR, RST, FCW, SB, SR, and BB had full access to all the data in the study and DAR had final responsibility for the decision to submit for publication.
Activation Scale BA competence was 55 (7·5) and mean Revised Cognitive Therapy Scale for CBT competence was 37·9 (10·9).

We found no evidence of inferiority of PHQ-9 score at 12 months in either the mITT (CBT 8·4 PHQ-9 points [SD 7·5]; BA 8·4 PHQ-9 points [7·0]; mean difference 0·1 [-1·3 to 1·5]* p=0·89 0·2 [-1·1 to 1·7]* p=0·80) or PP (CBT 7·9 PHQ-9 points [7·3]; 7·8 (6·5); 135 0·0 [-1·5 to 1·6]* p=0·99 0·0 [-1·6 to 1·6]* p=0·99) populations (table 2). The non-inferiority of BA to CBT was accepted for both the mITT and PP populations as the lower bound of the 95% CI (one-sided 97·5% CI) of the between-group mean difference lies within the non-inferiority margin of –1·9 PHQ-9 points (appendix). Although we initially planned to include therapist as a random-eff ects variable, given the low levels of observed clustering, we parsimoniously fi tted our models without therapist as a variable. We checked for no inference diff erence with and without inclusion of a random-eff ects therapist term. We ruled out superiority of CBT to BA as the lower bound of the 95% CI included zero for the mITT and PP populations. The inference of non-inferiority was robust to sensitivity analysis across different PP defi nitions. We found no evidence of a signifi cant between-group treatment interaction across the mITT or PP populations with the primary outcome at 12 months as stratifi ed by depression severity, antidepressants use [use or not], and recruitment site (appendix).

We found that BA was not diff erent from CBT in anxiety (Generalized Anxiety Disorder 7), depression status, and depression-free days and anxiety diagnoses (SCID) for either the mITT or PP populations using observed or imputed data at 12 months (table 2). Because of substantial missing 36-Item Short Form Survey data at baseline, we analysed these data adjusted for stratifi cation variables only. We found no diff erence in numbers of participants with at least one anxiety diagnosis: BA 43 (28%) of 153; CBT 43 (27%) of 161 (mITT population; χ² 0·08; p=0·78).

Between 61% and 70% of mITT and PP participants in both groups met criteria for recovery from depression or response to treatment at 12 months, with no diff erences in the proportions of patients in each group who recovered or responded (table 3). Using observed data for all outcomes, we found no evidence of a difference

<table>
<thead>
<tr>
<th>CBT</th>
<th>BA</th>
<th>Observed data only</th>
<th>Observed and imputed data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Between-group</td>
<td>p value</td>
</tr>
<tr>
<td>Primary outcome</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PHQ-9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>17·4 (4·8); 219</td>
<td>17·7 (4·8); 221</td>
<td>---</td>
</tr>
<tr>
<td>mITT</td>
<td>8·4 (7·5); 189</td>
<td>8·4 (7·0); 175</td>
<td>0·1 (-1·3 to 1·5)* p=0·89</td>
</tr>
<tr>
<td>PP</td>
<td>7·9 (7·3); 151</td>
<td>7·8 (6·5); 135</td>
<td>0·0 [-1·5 to 1·6]* p=0·99</td>
</tr>
<tr>
<td>Secondary outcomes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GAD-7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>12·6 (5·1); 219</td>
<td>12·7 (5·1); 221</td>
<td>---</td>
</tr>
<tr>
<td>mITT</td>
<td>6·3 (6·0); 176</td>
<td>6·4 (5·9); 161</td>
<td>-0·1 (-1·0 to 1·3)* p=0·82</td>
</tr>
<tr>
<td>PP</td>
<td>6·0 (5·8); 146</td>
<td>5·9 (5·5); 129</td>
<td>0·0 [-1·3 to 1·2]* p=0·95</td>
</tr>
<tr>
<td>SCID number of depression-free days</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>mITT</td>
<td>129 (58); 160</td>
<td>120 (56); 150</td>
<td>9 (-3 to 23)* p=0·13</td>
</tr>
<tr>
<td>PP</td>
<td>132 (55); 138</td>
<td>119 (55); 125</td>
<td>0·06</td>
</tr>
<tr>
<td>SF-36v2 PCS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>50·3 (13·3); 65</td>
<td>51·4 (13·1); 69</td>
<td>---</td>
</tr>
<tr>
<td>mITT</td>
<td>48·1 (12·2); 168</td>
<td>49·9 (11·6); 150</td>
<td>1·6 (-1·0 to 4·2)* p=0·22</td>
</tr>
<tr>
<td>PP</td>
<td>48·0 (12·2); 144</td>
<td>49·9 (12·0); 125</td>
<td>1·6 (-1·3 to 4·4)* p=0·28</td>
</tr>
<tr>
<td>SF-36v2 MCS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>23·2 (9·4); 65</td>
<td>22·5 (7·8); 69</td>
<td>---</td>
</tr>
<tr>
<td>mITT</td>
<td>41·7 (14·1); 168</td>
<td>41·6 (14·0); 150</td>
<td>0 (0·3 to 3·0)* p=0·99</td>
</tr>
<tr>
<td>PP</td>
<td>42·9 (13·6); 144</td>
<td>42·3 (13·3); 125</td>
<td>-0·5 (-3·7 to 2·7)* p=0·77</td>
</tr>
</tbody>
</table>

Data are mean (SD); n or mean (95% CI). CBT=cognitive behavioural therapy. BA=behavioural activation. PHQ-9=Patient Health Questionnaire 9. mITT=modifi ed intention to treat. PP=per protocol. GAD-7=Generalized Anxiety Disorder 7. SCID=Structured Clinical Interview for the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition. SF-36v2=36-Item Short Form Survey version 2. PCS=physical component summary. MCS=menta

Table 2: Primary and secondary outcomes at 12 months
between the CBT and BA groups over the period of the trial, as indicated by a non-significant time-by-treatment effect interaction, for both the mITT and PP populations (appendix). We found a small, negligible clustering of primary and secondary outcome scores at follow-up across therapists overall and within BA and CBT groups (intracluster correlation coefficient ≤0.04).

Two (1%) non-trial-related deaths (one [1%] multitrauma 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). Of the 440 participants recruited, 76 (17%) had missing primary outcome data at 12 month follow-up. The proportion of missing PHQ-9 data was higher in the BA than in the CBT group (46 [21%] vs 30 [14%]; odds ratio 1.6 [95% CI 1.0–2.7]; p=0.05). Imputation of data for primary and secondary outcomes at 12 months showed that in accordance with the observed data analysis, no difference existed between groups (tables 2, 3), supporting our conclusion of non-inferiority. The odds of missing PHQ-9 data were higher for patients with increased baseline severity of depression (PHQ ≥19, odds ratio 1.6 [95% CI 1.0–2.6]; p=0.05) and increasing age (in years) was associated with lower odds of missing PHQ-9 data (odds ratio 0.97 [0.96–0.99]; p=0.01). We found no evidence of an association between missingness and any other baseline characteristic (data not shown). Outcome assessors reported having been unmasked for 16 (4%) participants (five [2%] in the BA group and 11 [5%] in the CBT group; due to participants informing assessors of their treatment allocation).

For economic analyses, at 18 months, full service use data was available for 159 (90%) of 176 participants in the BA group and 168 (93%) of 180 participants in the CBT group. We found a significant difference in mean intervention costs between the two groups, but no differences in other categories of cost or in total cost (table 4). Mean health state utility scores according to EuroQol-5D-3L were slightly higher in the BA group than in the CBT group across the entire follow-up period, with resultant QALYs also higher for BA, but the QALY difference was not significant. Costs were lower and QALY outcomes better in the BA group than in the CBT group, generating an incremental cost-effectiveness ratio of −£6865. The scatterplot of bootstrapped cost and effectiveness pairs for BA versus CBT illustrates dominance of BA over CBT, with the point estimate and two-thirds of scatter points falling in the southeast quadrant of the cost-effectiveness plane, where BA replications are cheaper and more effective than are CBT ones (figure 2). The cost-effectiveness acceptability curve (appendix) showing the probability of BA being cost-effective compared with CBT does not fall below 75% and

<table>
<thead>
<tr>
<th>CBT</th>
<th>BA</th>
<th>Observed data only</th>
<th>Observed and imputed data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Odds ratio</td>
<td>p value</td>
<td>Odds ratio</td>
<td>p value</td>
</tr>
<tr>
<td>SCID depression</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>219/219 (100%)</td>
<td>221/221 (100%)</td>
<td>0.9 (0.9–1.1)</td>
</tr>
<tr>
<td>mITT</td>
<td>37/161 (23%)</td>
<td>31/154 (20%)</td>
<td>0.9 (0.5–1.1)</td>
</tr>
<tr>
<td>PP</td>
<td>30/141 (21%)</td>
<td>24/128 (19%)</td>
<td>0.9 (0.5–1.1)</td>
</tr>
<tr>
<td>Depression recovery*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mITT</td>
<td>124/189 (66%)</td>
<td>117/175 (66%)</td>
<td>0.9 (0.6–1.1)</td>
</tr>
<tr>
<td>PP</td>
<td>104/151 (69%)</td>
<td>97/128 (77%)</td>
<td>0.9 (0.6–1.1)</td>
</tr>
</tbody>
</table>

Data are n/N (%) or odds ratio (95% CI). CBT=cognitive behavioural therapy. BA=behavioural activation. SCID=Structured Clinical Interview for the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition. mITT=modified intention to treat. PP=per protocol. *Participants with a 50% reduction from baseline in Patient Health Questionnaire 9 score.

### Table 3: Depression status, recovery, and response at 12 months

<table>
<thead>
<tr>
<th>BA</th>
<th>CBT</th>
<th>Difference</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Costs per participant (£)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intervention</td>
<td>£7.48 (1.75)</td>
<td>£12.35 (2.10)</td>
<td>£4.87 (1.35)</td>
</tr>
<tr>
<td>Hospital</td>
<td>£860.32 (15.99)</td>
<td>£927.26 (19.15)</td>
<td>£67.94 (1.96)</td>
</tr>
<tr>
<td>Community health and social care</td>
<td>£644.36 (1.86)</td>
<td>£944.25 (17.26)</td>
<td>£300.89 (1.24)</td>
</tr>
<tr>
<td>Medication</td>
<td>£103.20 (1.92)</td>
<td>£117.64 (2.65)</td>
<td>£14.44 (1.24)</td>
</tr>
<tr>
<td>Total</td>
<td>£2596.62 (18.46)</td>
<td>£3250.74 (30.94)</td>
<td>£654.12 (1.24)</td>
</tr>
</tbody>
</table>

Data are mean (SD), n or mean difference (95% CI). BA=behavioural activation. CBT=cognitive behavioural therapy. EQ=EuroQol. QALY=quality-adjusted life-year.

### Table 4: Economic data at 18 months
is closer to 80% at NICE-preferred willingness\textsuperscript{44} to pay £20 000–30 000 per QALY.

In all sensitivity analyses, including complementary therapies and productivity losses, as well as analyses taking narrow intervention and mental health service perspectives, BA was significantly less costly than was CBT, so BA continues to have a higher probability of being cost-effective than does CBT at the NICE threshold (appendix).\textsuperscript{34} Imputation of missing data increased the difference in total cost (BA £1841.67; CBT £2282.40; difference −£440.73 [95% CI −1007.71 to 126.26]; p=0.13), but reduced the difference in QALYs (BA 1.22; CBT 1.19; difference 0.03 [0.06 to 0.11]; p=0.55), increasing the incremental cost-effectiveness ratio to −£16 951. The cost-effectiveness acceptability curve for the missing data analysis again supported the likelihood that BA is cost-effective compared with CBT (appendix).

**Discussion**

We found that BA for depression is not inferior to CBT in terms of reduction of depression symptoms and is more cost-effective than is CBT against commonly applied decision maker willingness to pay thresholds. We observed our results using both mITT and PP analyses, using a conservative non-inferiority margin. Our economic analyses 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 is CBT and that its simplicity renders BA suitable for delivery by junior MHWs with no professional training in psychological therapies.\textsuperscript{3}

This trial is the largest trial of BA to date and is one of the largest trials of psychological treatments 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 UK care settings delivered treatment, providing evidence of potential generalisability. We assessed therapy quality using independent raters and ensured that treatment in both arms was delivered to the standard recommended guidelines. Our levels of attrition and outcome loss to follow-up were low at 12 months and 18 months, similar to other trials in this area, but are still a limitation. Although participants in the per-protocol population attended similar numbers of sessions to those in other CBT trials,\textsuperscript{35} 35% of participants chose to not even attend a minimal number of sessions, a problem well known to routine psychological therapy services. This pragmatic trial done in routine environments means that we were unable to quantify or control for the contribution of antidepressant medicines to outcomes. However, most participants who were taking medication had been doing so for a considerable time before entering the trial, making it unlikely that our results were driven by pharmacological treatment. Given the nature of the intervention and comparator, we could not mask patients or the mental health workers or therapists who were delivering the interventions to treatment allocation, but we used self-reported outcome measures and robust outcome assessor-masking procedures to reduce researcher unmasking to less than 5%. Missing data for the primary outcome measure was substantial. However, our between-group inferences were robust to data imputation.

Our findings could have substantial implications for the scalability of psychological treatment for depression internationally\textsuperscript{36} given the greater availability and ease with which a BA workforce could be trained than could a CBT workforce. For many years, CBT has been the foremost psychological therapy recommended by therapists, researchers, and policy makers. Our results challenge this dominance. Although more work needs to be done than has been done so far to find ways to effectively treat the 20–23% of patients whose depression was unchanged by BA or CBT, our findings suggest that BA should be a front-line treatment for depression, with substantial potential to improve reach and access to psychological therapy globally.

Our results in both groups compare favourably with a meta-analysis\textsuperscript{1} of the effects of CBT that estimate proportions of patients with remissions of around 50%. Our cost-effectiveness analyses show the high probability that BA is cost-effective and affordable compared with CBT at standard willingness to pay thresholds. Our most striking finding is that BA leads to similar clinical outcomes for patients with depression, but at a financial saving to clinical providers of 21% compared with the costs of provision of CBT, with no compensatory use of other health-care services by patients.

Driving these savings is the fact that BA can be delivered by inexperienced MHWs with no professional training in psychological therapies, with no lesser effect than that of more highly trained and experienced psychological...
therapists giving patients CBT. 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.1 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-income and middle-income countries. Such countries might choose to investigate the training and employment of junior workers over expensive groups of psychological professionals. 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.

Contributors
DAR, DE, DM, RST, SB, PAF, SG, WK, HO’M, ERW, and KAW designed the study and were responsible for its conduct. SR, EF, and KEF were responsible for study and data collection management. RST, SB, FCW, and BB did data analysis. SDH and NR provided expert advice on clinical and patient and public involvement. All authors contributed to writing and editing of the manuscript.

Declaration of interests
All authors report grants from the National Institute for Health Research (NIHR) during the course of the study. DAR reports grants from the European Science Foundation. DAR and RST have received funding support from NIHR Collaborations for Leadership in Applied Health Research and Care and report NIHR panel memberships. WK reports fees from book royalties.

Acknowledgments
This report is independent research funded by the UK National Institute for Health Research (NIHR) Health Technology Assessment Programme. The views expressed in this publication are those of the authors and not necessarily of the NIHR or UK Department of Health. We would like to thank all participants, National Health Service services, mental health workers, therapists, and general practitioners involved in the study and acknowledge the vital contributions of study researchers and administrators in Devon, Durham, and Leeds, the Peninsula Clinical Trials Unit, and the NIHR Clinical Research Network.

References