Psychological Medicine (2008), **38**, 611–623. © 2007 Cambridge University Press
doi:10.1017/S0033291707001614 Printed in the United Kingdom

REVIEW ARTICLE

A meta-analysis of randomized trials of behavioural treatment of depression

D. Ekers[1]*, D. Richards[2] and S. Gilbody[2]

[1] *Tees Esk & Wear Valleys NHS Trust/University of York Department of Health Sciences, The Health Centre, Newcastle Road, Chester le Street, Co. Durham, UK*
[2] *Department of Health Sciences, Seebohm Rowntree Building, University of York, York, UK*

Background. Depression is a common, disabling condition for which psychological treatments, in particular cognitive behavioural therapies are recommended. Promising results in recent randomized trials have renewed interest in behavioural therapy. This systematic review sought to identify all randomized trials of behavioural therapy for depression, determine the effect of such interventions and examine any moderators of such effect.

Method. Randomized trials of behavioural treatments of depression *versus* controls or other psychotherapies were identified using electronic database searches, previous reviews and reference lists. Data on symptom-level, recovery/dropout rate and study-level moderators (study quality, number of sessions, severity and level of training) were extracted and analysed using meta-analysis and meta-regression respectively.

Results. Seventeen randomized controlled trials including 1109 subjects were included in this meta-analysis. A random-effects meta-analysis of symptom-level post-treatment showed behavioural therapies were superior to controls [standardized mean difference (SMD) −0.70, 95% CI −1.00 to −0.39, $k=12$, $n=459$], brief psychotherapy (SMD −0.56, 95% CI −1.0 to −0.12, $k=3$, $n=166$), supportive therapy (SMD −0.75, 95% CI −1.37 to −0.14, $k=2$, $n=45$) and equal to cognitive behavioural therapy (SMD 0.08, 95% CI −0.14 to 0.30, $k=12$, $n=476$).

Conclusions. The results in this study indicate behavioural therapy is an effective treatment for depression with outcomes equal to that of the current recommended psychological intervention. Future research needs to address issues of parsimony of such interventions.

Received 30 March 2007; Revised 25 July 2007; Accepted 31 July 2007; First published online 1 October 2007

Key words: Behavioural therapy, cognitive behavioural therapy, depression, meta-analysis, psychotherapy.

Introduction

Depression causes substantial disability, is set to become the second largest cause of disease burden by 2020 (WHO, 2001), affects between 5% and 10% of the population and is the third most common reason for primary-care consultation (Singleton *et al*. 2001). It is associated with significant distress, impairment of functioning, disturbance to interpersonal relationships and an increased risk of suicide (Hirschfeld *et al*. 1997). Psychological treatments, particularly cognitive behavioural therapy (CBT) are recommended to treat depression (Hollon *et al*. 2002; NICE, 2004), however, less than 10% of those affected receive such treatment (Singleton *et al*. 2001).

* Address for correspondence: D. Ekers, Tees Esk & Wear Valleys NHS Trust/Department of Health Sciences, The Health Centre, Newcastle Road, Chester le Street, Co. Durham DH3 3UR, UK.
(Email: David.Ekers@cddps.nhs.uk)

CBT combines both behavioural and cognitive techniques in each treatment programme. The standard approach is Beck's cognitive therapy (Beck, 1976) using both behavioural and cognitive techniques to identify, question and modify maladaptive thought processes, life rules and core beliefs.

However, recent research has suggested that pure behavioural models utilizing an operant conditioning formulation to develop a structured daily action plan may be as effective as full cognitive therapy (CT) (Jacobson & Gortner, 2000, Jacobson *et al*. 2001). Ferster (1973) pioneered the early incorporation of learning theory to the treatment of depression in the 1970s followed by the establishment of the 'coping with depression' intervention (Lewinsohn & Graf, 1973). With the development of cognitive models, behavioural interventions lost popularity, until recent renewed interest led to research reminding us of their potential (Jacobson & Gortner, 2000). The optimum combination of behavioural and cognitive techniques within CBT is unknown (Jacobson *et al*. 1996).

Behavioural therapy (BT) may provide a more parsimonious treatment option as it may be simpler to deliver (Jacobson *et al*. 1996). If similar health outcomes could be achieved with such a lesser 'dose' of psychotherapy, service and training procedures could be radically overhauled. Narrative reviews conducted by advocates for behavioural approaches suggest positive outcomes (Martell *et al*. 2001, Hopko *et al*. 2003), however, such reviews are prone to bias. Alternatively, systematic reviews of psychotherapy for depression have looked at behavioural interventions in the context of considering the effect of other psychological approaches (Dobson, 1989; Gloaguen *et al*. 1998; Churchill *et al*. 2001).

Therefore, we conducted a systematic review of randomized controlled trials of behavioural interventions for depression compared to other psychological approaches and controls. We explored effectiveness in terms of depressive symptoms, dropout and recovery rates.

Method

Identification of suitable studies

We searched a range of databases from inception to January 2006 (Medline, EMBASE, PsycINFO, Cochrane Library DARE, CINAHL, AMED and the British Nursing Index), incorporating randomized controlled trial filters. We reviewed reference lists of identified studies to find additional trials. Two authors (D.E. and D.R.) considered abstracts and screened the full text of selected studies for relevance.

Inclusion criteria

We included all available randomized controlled trials in any language to reduce the potential for publication bias (Khan & Kleijnen, 2002). Studies included participants who were adults (aged $\geqslant$16 years), treated in community or in-patient settings with a primary diagnosis of depression. We excluded studies including participants with psychosis or bipolar disorder, substance misuse problems, cognitive impairment or without depression as primary diagnosis. We included trials of individual time-limited behaviourally orientated psychotherapeutic approaches to the treatment of depression with an alternative psychotherapy, control with confirmation of randomized allocation.

BT

We included trials in the behavioural intervention group if the treatment was based upon the rescheduling of activities to reintroduce positive reinforcement and reduce avoidance. Such interventions manipulate the behavioural consequence of a trigger (environmental or cognitive) rather than directly interpret or restructure cognitions.

Comparators

Treatment as usual/control. A range of standard treatments or non-treatment options (waiting list, usual general practitioner treatment, inert control conditions) delivered to the patient in the absence of any 'active' psychotherapy.

CBT/CT. Interventions that directly identified, questioned and modified cognitive responses to situations and their emotional consequences. We included any intervention conceptualized as an intervention to directly challenge thinking including 'thought catching' and 'challenging' through diary-keeping or behavioural experiments.

Brief psychotherapy. Approaches that focused on developing insight and subsequent character development through interpersonal relationships with the therapist, including brief interpersonal therapy (IPT; Klerman *et al*. 1984) or brief psychodynamic therapy (Luborsky *et al*. 1995).

Supportive counseling. We included any approach which focused upon the therapist's use of core relationship conditions (Rogers, 1961) to develop self-awareness by the patient.

We excluded marital, couple or group therapy as the change in therapist contact coupled with other group-member interaction would introduce substantial clinical heterogeneity and was outside the aims of this review.

Outcome measures

Our primary outcome measure was depression symptom-level self-rated [e.g. Beck Depression Inventory (BDI); Beck *et al*. 1961] or clinician-rated [e.g. Hamilton Depression Rating Scale (HAMD); Hamilton 1960], presented by means and standard deviations (continuous data) or clinical improvement/non-clinical improvement (dichotomous data). As psychotherapy trials often present multiple symptom measures we adopted an algorithm so that validated self-report measures took precedence over clinician-rated measures and performed sensitivity

analysis to explore the impact of this approach. We entered recovery and dropout rates as dichotomous data, dropout being viewed as a proxy for acceptability.

Quality assessment

Two authors (D.E. and D.R.) rated study quality using criteria to explore bias (Khan *et al*. 2002). Other than concealment of allocation, clear guidance on aspects of study quality that directly influence outcomes is unclear (Jadad *et al*. 1996; Schultz & Grimes, 2002). We assessed studies against two standards each relating to selection, measurement, performance and attrition bias resulting in an overall score of between 1 and 8. Disagreements regarding study quality were dealt with through discussion.

Data extraction and synthesis

We extracted data from each trial at post-treatment and follow-up (6 months or nearest available dataset). We synthesized data using the Cochrane collaboration RevMan program (Cochrane Collaboration, 2003). We sought missing data from study authors by email. We imputed missing standard deviation (S.D.) scores from other relevant studies where these data were not available following the above procedure (Furukawa *et al*. 2006).

Data pooling

We combined continuous data to estimate the standardized mean difference (SMD) across trials to facilitate analysis of the same outcome (depression symptom level) using different scales as a standardized unit (SMD). Where studies included two comparisons under the same category (i.e. CT and CBT) we entered these comparisons separately but halved numbers in the behavioural arm to avoid double counting and inaccurate weighting of trials. Where studies presented results using subcategories (e.g. high/low depression severity), we entered data as two separate trials, provided that stratification occurred prior to randomization. We assigned effect sizes according to the standard convention where the SMD is small (0–0.32), medium (0.33–0.55) and large ($\geqslant$0.56) (Lipsey & Wilson, 1993). We present dichotomous data for dropout and recovery rate as odds ratios (OR), which demonstrates the chance of an event (improvement or dropout) in the intervention group compared to the comparison group. We present pooled data with 95% confidence intervals (CI) using a random-effects model (Sutton *et al*. 1998) taking into account both within- and between-study variance. We consider such a model as appropriate based upon anticipated heterogeneity for this review (number of sessions, therapy approaches and setting, etc.).

Exploration of heterogeneity

We measured statistical heterogeneity using the I^2 statistic for statistical variation across studies (Higgins *et al*. 2003); values of 25% are low, 50% moderate and 75% high.

Three sources of clinical and statistical heterogeneity were identified *a priori*: (1) baseline severity of depression; (2) training level of the therapist (graduate *versus* postgraduate/experienced therapist qualification); (3) number of sessions. We considered study quality as a source of potential heterogeneity, by assessing the impact of lower quality studies on overall outcomes; using a cut-point of 6 on the 8-point quality scale.

We explored the impact of these sources of heterogeneity using sensitivity analyses and meta-regression (Thompson & Higgins, 2002). We analysed outcomes using meta-regression, specifying sources of heterogeneity as predictive covariates. We used a permutation test (using 1000 Monte-Carlo simulations) to calculate p values, and to reduce spurious false-positive findings (Higgins & Thompson, 2004). The amount of heterogeneity explained by predictive covariates was examined by reductions in the I^2 inconsistency statistic within our model. Analyses were conducted using the METAN and METAREG commands in Stata 8 (Stata Corporation, 2003).

The possibility of publication bias was assessed through a Begg funnel plot graph (Begg, 1994) and testing for asymmetry using the Egger weighted regression test (Egger *et al*. 1997) where the intercept is 0 if no bias is present.

Results

Searches conducted between December 2005 and February 2006 identified 3353 studies (see Fig. A1 for study flow chart; available in online Appendix). We identified 20 randomized controlled trials (Table 1), three of which were excluded from the meta-analysis [18–20] due to insufficient reported data. (*Note*: throughout the following sections numbers within square brackets refer to the Study numbers listed in Table 1.) We meta-analysed the remaining studies which included 1109 subjects (Table 2).

Table 1. *List of studies included in review*

Study no. (first-named author and year)	Sample/setting Mean age [S.D. (range)] Sex (% female)	Interventions (*n* in cell)	Depression level at baseline	Concurrent pharmacology	Therapist level Session number (duration)
[1] Taylor (1977)	University students 22.4 (2.6) 71	Behavioural (7) Cognitive (7) Cognitive behavioural (7) Wait list (7)	Mild/moderate (21.2 BDI)	No	Graduate student experience as counsellor 6 (40 min)
[2] McLean (1979)	Community out-patient 39.2 (10.9) 72	Behavioural (42) Brief psychotherapy (44) Drug therapy (49) Relaxation (43)	Within or beyond moderate depression range 2 out of 3 measures used at baseline	No (other than DT arm)	Licensed psychologists, physicians or psychiatrists. At least 2 years of experience as therapist 10 (1 h) not drug therapy
[3] Gallagher (1982)	Older adult community 67.76 (6) 76	Behavioural therapy (10) Cognitive therapy (10) Psychotherapy (10)	RDC Criteria MDD	No	Advanced Ph.D. or post-doctoral therapists experience in modality 16 (90 min)
[4] Maldonado Lopez (1982)	Community out-patient N.A. N.A.	Behavioural (8) Cognitive (8) Drug therapy (8)	Psychiatrist Diagnosis reactive depressive disorder	No (other than DT arm)	Psychology Dept, level of training not reported 10 (1 h)
[5] Wilson (1982)	General population media announcements 38.8 (20–55) 66	Drug therapy & Behavioural (12) Drug therapy & Relaxation (10) Drug therapy & Minimal contact (10) Placebo & Behavioural (9) Placebo & Relaxation (11) Placebo & Minimal contact (12)	BDI > 19	In DT arm	Graduate psychologist 7 (1 h) 2 (1 h) in min contact arm
[6] Wilson (1983)	General population media announcements 39.5 (20–58) 80	Behavioural (8) Cognitive (8) Wait list (9)	BDI > 17 (moderate depression)	Yes (5 subjects in trial)	Not clear, University Psychology Clinic 8 (1 h)
[7] Cole (1983)	Community out-patient veterans 56 (24–71)	Behavioural (15) Treatment as usual (15)	Psychiatrist Diagnosis major depression. BDI > 24	Yes if stable	Doctoral clinical psychology student 7 (1 h)
[8] Maldonado Lopez (1984)	Community out-patients N.A. N.A.	Behavioural & Pharmacology (8) Cognitive & Pharmacology (8) Pharmacology (8)	Psychiatrist Diagnosis reactive depressive disorder	All subjects	Psychology Dept, level of training not reported 10 (1 h)
[9] Skinner (1984)	Community volunteers 20–61 (34) 67.5	Behavioural (8) Cognitive(7) Control (9)	BDI > 12	Yes	Doctoral clinical psychology student 5 (1 h)
[10] McNamara (1986)	University students 23 (19–31) 73	Behavioural (10) Cognitive (10) Cognitive behavioural (10) Supportive (10)	BDI > 17 HAMD > 20	Not reported	Doctoral interns in clinical psychology/ masters-level social worker 8 (50 min) (10 sessions in CBT arm)

[11] Thompson (1987)	Older adults community 67.07 (5.8) 67	Behavioural (25) Cognitive (27) Psychotherapy (24) Delayed (19)	RDC MDD	If stable dose for 3 months	Doctoral-level psychologists plus 1 year specialized therapy training 16–20 (duration of each session not reported)
[12] Scogin (1989)	Older adults community 68.3 (6.7) 85	Behavioural bibliotherapy (23) Cognitive bibliotherapy (22) Delayed (22)	>9 on HAMD	If stabilized prior to trial	N.A. as bibliotherapy was main intervention 4 (5 min) phone contacts to support exercises
[13] Jacobson (1996)	Community (80% HMO, 20% volunteer) 38 (not reported) 72	Behavioural (56) Thought challenging (43) Full cognitive (50)	Major depression (DSM-IV) >19 BDI	No	Experienced therapists (mean 9.5 years CT practice) 20 sessions (N.A.)
[14] McKendree Smith (1998)	Community volunteer 44.88 (13.17) 75	Behavioural bibliotherapy (13) Cogntive bibliotherapy (13) Delayed control (14)	Mild-moderate depression	If stabilized for 3 months	N.A. as Bibliotherapy main intervention 8 (10 min)
[15] Hopko (2003*a*)	In-patients 30.5 (9) 36	Behavioural (10) Supportive (15)	Principle diagnosis of major depression	Yes all patients	Not clear 6 (20 min)
[16] Dimidjian (2006)	Community 39.9 (10.97) 66	Behavioural (43) Cognitive (45) Pharmacology/Placebo (153)	Major depression (DSM-IV) >19 BDI	Only in ADM arm	BA-licensed psychologists/social worker (7 years practice. CT-licensed psychologists with CT training 24 (50 min)
[17] Cullen (2006)	Community 38.48 (12.69) 32	Behavioural (13) Wait list (12)	MDD (Mean BDI) 30.96 (5.90)	Yes if stable >6 weeks	Previous exp. in CT of depression plus 12 h training in BA 10 (50)
Studies not included in meta-analysis					
[18] Padfield (1976)	Community female rural low socio-economic status 21–56 100	Behavioural (12) Supportive (12)	Moderately depressed (diagnostic tool not clear)	No	Counsellor (experience not clear) 12 (N.A.)
[19] Zeiss (1979)	Community 33.9 (19–68) N.A.	Behavioural (22) Cognitive (22) Interpersonal (22)	Classed as depressed using Minnesota Multiphasic Personality Inventory & Grinkler Interview Rating	Not clear	Graduate students in clinical psychology & counselling psychologists (masters level). At least 1 year experience 12 (N.A.)
[20] Gardner (1981)	Community 19–65 77%	Behavioural (8) Cognitive (8)	Mild depression (BDI)	Not clear	Not clear N.A.

ADM, Antidepressant medication; BT, behavioural therapy; BDI, Beck Depression Inventory; CBT, cognitive behavioural therapy; CT, cognitive therapy; DT, drug therapy; HAMD, Hamilton Depression Rating Scale; HMO, health maintenance organization; MDD, major depressive disorder; n.a., not available; RDC, research diagnostic criteria.

Table 2. *Meta-analyses of studies examining the effects of behavioural therapy*

Comparison	No. of studies	No. of subjects	SMD	95% CI	*p*	I^2
BT *versus* Control/TAU						
Symptom level	12	459	−0.70	−1.00 to −0.39	<0.001	55.1%
Dropout[a]	3	119	0.58	0.28 to 1.20	0.86	0%
Recovery rate[a]	3	167	4.18	1.14 to 15.28	0.03	52.6%
BT *versus* CT/CBT						
Symptom-level post-treatment	12	476	0.08	−0.14 to 0.30	0.46	21.1%
Symptom-level follow-up	8	271	0.25	−0.21 to 0.70	0.28	60.2%
Dropout[a]	8	436	1.17	0.57 to 2.41	1.17	32.4%
Recovery rate[a]	5	346	0.92	0.59 to 1.44	0.92	0%
BT *versus* Brief psychotherapy						
Symptom-level post-treatment	3	166	−0.56	−1.0 to −0.12	0.01	43.4%
Symptom-level follow-up	2	96	−0.50	−0.90 to −0.09	0.02	0%
Dropout[a]	3	166	0.94	0.22 to 3.96	0.11	54.1%
Recovery rate[a]	3	164	2.37	1.23 to 4.57	0.01	0%
BT *versus* Supportive therapy						
Symptom-level post-treatment	2	45	−0.75	−1.37 to −0.14	0.02	0%

BT, Behavioural therapy; CBT, cognitive behavioural therapy; CT, cognitive therapy; CI, confidence interval; SMD, Standardized mean difference; TAU, treatment as usual.

[a] Indicates odds ratio.

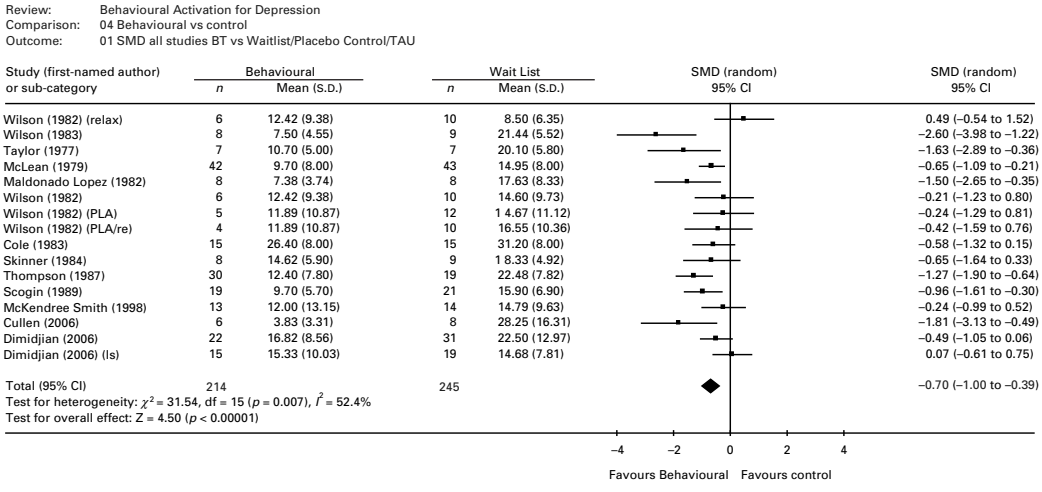

Fig. 1. Behavioural therapy (BT) *versus* wait list/control/placebo symptom-level post-treatment.

Comparison 1: Behavioural interventions versus *waiting list/placebo control*

Scope

Twelve studies with a total of 459 patients contributed data to this analysis [1, 2, 4–7, 9, 11, 12, 14, 16, 17]. Participants were taken from adult community sources consisting of out-patients [2, 4, 6, 7, 11, 12, 16, 17], volunteers [5, 8, 14] and students [1], two studies used older adults [11, 12]. Control interventions consisted of delayed treatment [1, 3, 9, 11, 12, 14, 16, 17], treatment as usual [4, 5, 7] and relaxation [2, 5]. All comparisons were taken immediately after intervention. Interventions ranged from supported bibliotherapy [12, 14], brief therapy with six 40-min sessions [1] to 24 50-min sessions [16]. Facilitators were advanced graduate psychology/therapy students in five studies [1, 5, 6, 7, 9], experienced psychotherapists in four studies [2, 11, 16, 17] and unclear in one study [4]. Depression symptom level was assessed using

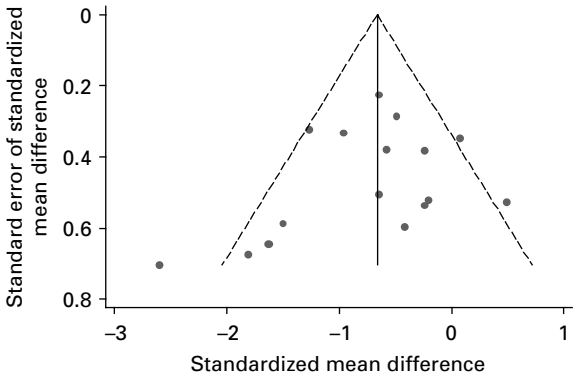

Fig. 2. Begg funnel plot symptom level: behavioural therapy *versus* control interventions/usual care.

either BDI self-report measure [1, 2, 4, 5, 7, 9, 17] or the HAMD assessor rating scale [12], or both [6, 11, 14, 16]. Recovery was defined by clinical interview in one study [11] and by BDI score in two studies [2, 14].

Outcome 1: Depression symptom level post-treatment

The effect of behavioural interventions against control interventions was large with a pooled SMD of −0.70 (95% CI −1.00 to −0.39), demonstrating a highly significant difference in symptom-level scores favouring the behavioural group ($p<0.001$) (Fig. 1). There was no evidence of publication bias for this outcome (Eggers test −1.04; 95% CI −3.39 to 1.29, $p=0.35$), a funnel plot showed no evidence of asymmetry (Fig. 2).

Heterogeneity and sensitivity analysis

Variation in effect size (I^2) attributable to heterogeneity was 55.1%. Effect size was not significantly related to the level of baseline severity (meta-regression β-coefficient 0.04, 95% CI −0.04 to 0.12; $I^2=54\%$, $p=0.28$) (Fig. A2 online). Quality assessment indicated seven studies fell below our quality threshold [1, 4–7, 9, 14], and the pooled SMD was not affected by study quality (meta-regression $SMD_{low\ quality}=-0.67$; $SMD_{higher\ quality}=-0.75$, $p_{difference}=0.77$). Behavioural therapists with graduate and postgraduate qualifications produced similar effect sizes (meta-regression $SMD_{graduate}=-0.82$; $SMD_{postgraduate}=-0.59$, $p_{difference}=0.61$; $I^2=59\%$). There was no clear relationship between effect size and number of sessions (meta-regression β-coefficient 0.03; 95%CI −0.03 to 0.09; $I^2=0.49$, $p=0.27$) (Fig. A3 online). Prioritizing clinician-rated assessment in precedence over self-rated where possible made no significant difference to overall effect size (SMD −0.68, 95% CI −0.98 to −0.38).

Outcome 2: Dropout rate

Three studies contributed data to this analysis [2, 14, 16] on a total of 119 subjects with an average dropout rate of 19.17%. We found no difference between rates of dropout between intervention and control (OR 0.58, 95% CI 0.28–1.20, $p=0.86$).

Heterogeneity and sensitivity analysis

Variation in effect size (I^2) attributable to heterogeneity was 0%. There were insufficient studies and negligible heterogeneity to explore the impact of our *a priori* sources of clinical heterogeneity.

Outcome 3: Recovery rate

Three studies contributed data to this analysis [2, 11, 14] on a total of 167 subjects. There were greater rates of recovery in the behavioural intervention group (BT 52%, control 21.05%) with an odds ratio of 4.18 (95% CI 1.14–15.28, $p=0.03$). There were insufficient studies to test for publication bias for this outcome.

Heterogeneity and sensitivity analysis

Variation in effect size attributable to heterogeneity (I^2) was 52.6%. Low-quality studies [14] were excluded in a sensitivity analysis resulting in an odds ratio of 8.56 (95% CI 0.40–182.63, $p=0.04$) with an I^2 statistic of 76.4%. There were insufficient studies to explore the underlying causes of this heterogeneity further.

Comparison 2: BT* versus *CT/CBT

Scope

Twelve studies with a total of 476 patients contributed data to this analysis [1, 3, 4, 6, 8–14, 16]. Participants were taken from adult community sources consisting of out-patients [3, 4, 8, 11–13, 16], volunteers [6, 9, 14] and students [1, 10], with three studies using older adults [3, 11, 12]. Interventions ranged from supported bibliotherapy [12, 14], brief therapy with six 40-min sessions [1] to 24 50-min sessions [16]. Therapy was facilitated by advanced graduate psychology/therapy students in four studies [1, 6, 9, 10], experienced psychotherapists in four studies [3, 11, 13, 16] and was unclear in two studies [4, 7]. Depression symptom level was assessed using either the BDI self-report measure [1, 4, 8–10] or the HAMD assessor rating scale [12], or both [3, 6, 11, 13, 14, 16]. Recovery was defined by diagnostic interview in two studies [3, 11] and by BDI score in three studies [10, 13, 16].

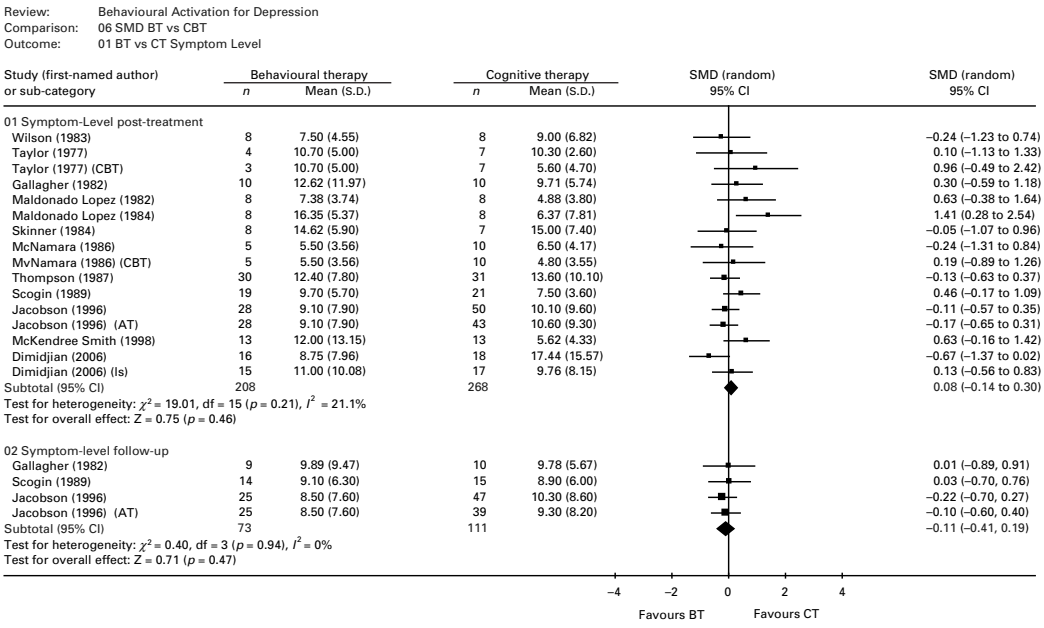

Fig. 3. Behavioural therapy (BT) *versus* cognitive behavioural therapy (CBT) symptom-level post-treatment and follow-up.

Outcome 1: Depression symptom level post-treatment

No difference in effect between behavioural interventions and CBT/CT was identified with a pooled SMD of 0.08 (95% CI −0.14 to 0.30, $p=0.46$) (see Fig. 3). There was no evidence of publication bias for this outcome using Egger's test [intercept (0 if unbiased) = 1.07; 95% CI −0.23 to 2.38, $p=0.10$], and a funnel plot showed no evidence of asymmetry.

Heterogeneity and sensitivity analysis

Variation in effect size (I^2) attributable to heterogeneity was 21.1%. Seven studies fell below our quality threshold [1, 4, 6, 7, 9, 10, 14] and the pooled SMD was not significantly affected by study quality (meta-regression $SMD_{low\ quality} = +0.23$; $SMD_{higher\ quality} = -0.13$, $p_{difference}=0.12$; $I^2=0\%$). Comparative effectiveness of BT *versus* CT/CBT varied according to baseline severity of depression, BT demonstrating a greater level of effectiveness at more severe levels of depression (meta-regression β-coefficient −0.05, 95%CI −0.10 to −0.01; $I^2=0\%$, $p=0.04$) (Fig. A4 online).

Graduate-level behavioural therapists produced slightly worse results compared to those with post-graduate qualifications in comparison to CBT, although this did not reach significance (meta-regression $SMD_{graduate}=0.28$; $SMD_{postgraduate} = -0.135$, $p_{difference}=0.11$; $I^2=0\%$). There was no clear relationship between effect size and number of sessions (meta-regression β-coefficient −0.025, 95% CI −0.056 to 0.006; $I^2=0.08$, $p=0.11$). Prioritizing clinician-rated assessment in precedence over self-rated where possible made no significant difference in overall effect size (SMD 0.09, 95% CI −0.12 to 0.29).

Outcome 2: Depression symptom level at follow-up

Eight studies contributed data to this analysis [1, 3, 4, 6, 8, 10, 12, 13] on a total of 271 subjects with an average follow-up period of 4 months. Overall there was no difference in effect of BT compared to CBT/CT with a pooled SMD of 0.25 (95% CI −0.21 to 0.70, $p=0.28$) (Fig. 3).

Heterogeneity and sensitivity analysis

Variation in effect size (I^2) attributable to heterogeneity was 60.2%. After exclusion of low- quality studies [1, 4, 6, 8, 10] and those with follow-up of <3 months [1, 6] in a sensitivity analysis the pooled SMD was −0.11 (95% CI −0.41 to 0.19, $p=0.47$). There were insufficient studies to explore the underlying causes of this heterogeneity further.

Outcome 3: Dropout rate

Eight studies contributed data to this analysis [1, 3, 6, 11–14, 16] on a total of 436 subjects with an average dropout rate of 15.36%. We found no difference in rates of dropout with an odds ratio of 1.17 (95% CI 0.57–2.41, $p=0.67$).

Heterogeneity and sensitivity analysis

Variation in effect size attributable to heterogeneity (I^2) was 32.4%. Low-quality studies [1, 6, 14] were excluded in a sensitivity analysis resulting in an odds ratio of 1.47 (95% CI 0.60–3.61, $p=0.40$) with an I^2 statistic of 42.9%. There were insufficient studies to explore the underlying causes of this heterogeneity further.

Outcome 4: Recovery rate

Five studies contributed data to this analysis [3, 10, 11, 13, 16] on a total of 346 subjects. We found a pooled recovery rate of 55% with no difference between the two treatment approaches (OR 0.92, 95% CI 0.59–1.44, $p=0.72$).

Heterogeneity and sensitivity analysis

Variation in effect size attributable to heterogeneity (I^2) was 0%. Low-quality studies [8] were excluded in a sensitivity analysis resulting in an odds ratio of 0.93 (95% CI 0.59–1.47, $p=0.77$) with an I^2 statistic of 0%.

Comparison 3: Behavioural interventions versus *brief psychotherapy*

Scope

Three studies with a total of 166 patients contributed data to this analysis [2, 3, 11]. Participants were from adult out-patient community sources, two studies using older adults [3, 11]. Brief psychotherapy interventions were based upon a psychodynamic model in all studies. Interventions ranged from 10 to 20 sessions, all studies used experienced therapists. Studies assessed depression symptom level using the BDI alone [2] or both BDI and HAMD [3, 11]. Two studies assessed depression at intake using structured clinical interviews [3, 11], the third using cut-off points from validated self-report measures [2]. Recovery was defined by clinical interview in two studies [3, 11] and by BDI score in one study [2].

Outcome 1: Depression symptom post-treatment

The positive effect of BT against brief psychotherapy was large with a pooled SMD of −0.56 (95% CI −1.0 to −0.12, $p=0.01$). There were insufficient studies to test for publication bias.

Heterogeneity and sensitivity analysis

Variation in effect size attributable to heterogeneity (I^2) was 43.4%. All studies were above the quality threshold, hence we performed no sensitivity analyses. There were insufficient studies to explore the underlying causes of this heterogeneity further. Prioritizing clinician- rated assessment in precedence to self-rated assessment where possible made no difference in overall effect size (SMD −0.52, 95% CI −1.01 to −0.03).

Outcome 2: Depression symptom level follow-up

Two studies contributed data to this analysis [2, 3] on a total of 96 subjects with an average follow-up period of 4.5 months. The positive effect of behavioural interventions against brief psychotherapy was medium with a SMD of −0.50 (95% CI −0.90 to −0.09, $p=0.02$).

Heterogeneity and sensitivity analysis

Variation in effect size attributable to heterogeneity (I^2) was 0%. Both studies collected follow-up beyond the 3-month point and were above the quality threshold so we performed no sensitivity analyses.

Outcome 3: Dropout

Three studies contributed data to this analysis [2, 3, 11] on a total of 166 subjects with an average dropout rate of 14.45% across studies. No difference in dropout was observed with an odd ratio of 0.94 (95% CI 0.22–3.96, $p=0.11$). There were insufficient studies to test for publication bias.

Heterogeneity and sensitivity analysis

Variation in odds ratio attributable to heterogeneity (I^2) was 54.1%. All studies were above the quality threshold so no sensitivity analysis was performed.

Outcome 4: Recovery rate

Three trials contributed data to this analysis [2, 3, 11] on a total of 164 subjects (note two subjects deceased). Greater rates of recovery were observed in BT (56.79%) compared to brief psychotherapy (36.14%) with an odds ratio of 2.37 (95% CI 1.23–4.57, $p=0.01$). There were insufficient trials to test for publication bias.

Heterogeneity and sensitivity analysis

Variation in odds ratio attributable to heterogeneity (I^2) was 0%. All studies were above the quality threshold so no sensitivity analysis was performed.

Comparison 4: Behavioural interventions versus *supportive therapy*

Scope

Two studies with 45 subjects contributed data to this analysis [10, 15]. Participants were university students [10] and in-patients [15].

Interventions ranged from six 20-min sessions [10] to eight, 50-min sessions [15], delivered by doctoral clinical psychology students [10] or a clinical psychologist [15]. Both studies measured depression symptom levels by self-report measures (BDI), with one [10] using HAMD also. Depression at baseline was assessed by self-report measures [10] or clinical interview [15].

Outcome 1: Depression symptom level post-treatment

The positive effect of BT against supportive therapy was large (SMD −0.75, 95% CI −1.37 to −0.14, $p=0.02$). There were insufficient studies to test for publication bias.

Heterogeneity and sensitivity analysis

The variation in effect size attributable to heterogeneity (I^2) was 0%. Both studies fell below the quality threshold therefore no sensitivity analysis was performed. Insufficient data were available in this comparison for further analysis.

Discussion

We found clear evidence that BT is an effective treatment for depression. It provides superior outcomes to control, supportive counselling and brief psychotherapy. BT and CBT provided equivalent results with no statistically significant differences in post-treatment and follow-up symptom levels, in recovery rate or dropouts.

The BT trials were variable in design and delivery. To some degree, we have been able to utilize this variability to explore factors relating to magnitude of effectiveness. Such meta-regression analysis makes observational associations and is exploratory in nature and as such loses the power of causal inference (Higgins & Thompson, 2004). We considered such an approach viable and efficient in this review as the alternative of planning large-scale prospective trials with many arms is costly and time consuming. Sufficient data for this analysis were available only where BT was compared to controls or CT/CBT post-treatment on symptom level. Our meta-regression found that compared to controls, baseline severity, length of treatment and level of qualification were not related to BT effect although there is a positive relationship between greater baseline severity and BT efficacy compared to CT/CBT. Such findings provide direction in the development of BT for future research. They indicated that further exploration is needed into length of treatment and skill level required for optimum BT delivery. Our review identified a number of trials directly comparing BT with drug therapy; this was not included as an *a priori* comparator. Such a comparison would be a useful addition in any future review as BT would appear equivalent, if not superior, to pharmacology in the included studies.

Our meta-analysis complements and concurs with other publications that include behavioural interventions as part of wider CBT reviews (Dobson, 1989; Gloaguen *et al*. 1998; Churchill *et al*. 2001), or focus on activation alone (Cuijpers *et al*. 2007). In contrast to these previous reviews we chose to focus on individual rather than group interventions, and included dropout and recovery rate analyses. Our review includes more studies than previous reviews due to our broader inclusion criteria and the inclusion of recent and unpublished data. The studies drew patients from a range of settings such as in-patient, psychiatric out-patient and volunteer cohorts in adult, older adult and student settings. Interventions varied considerably across studies from supported self-help using minimal therapist contact to full psychotherapy. The quality of included trials varied considerably, with some of low quality delivering results that deviated considerably from the overall picture [4, 8]. We attempted to account for this by the use of sensitivity analysis, random-effects modelling and meta-regression of *a priori* variables. Interpretation of our results must be made with such factors in mind. Caution must also be exercised in interpreting the comparisons of behavioural interventions with brief psychotherapy and supportive therapy due to the low numbers of studies and small sample sizes.

Of particular interest is the observed equivalence between behavioural interventions and the CBT/CT strongly recommended in guidelines (e.g. NICE, 2004). In addition to similar levels of mean symptom improvement, we observed no difference in recovery or dropout. These combined findings indicate that behavioural interventions are as effective and acceptable as CBT/CT. Such findings partially endorse the BT parsimony hypothesis advanced by Jacobson and colleagues (Jacobson *et al*. 1996, 2001). They question the utility of adding 'complex' cognitive techniques to simpler behavioural interventions to improve clinical outcome. One of the attractions of behavioural interventions is that they may lend themselves to shorter training of less-qualified individuals, thus assisting

the current scarcity of therapist availability and overwhelming demand (Centre for Economic Performance's Mental Health Policy Group, 2006). We found no direct evidence in this review to support such an assumption, as we found no studies using non-psychology- or non-psychotherapy-trained individuals delivering BT. However, when we examined the impact on level of training of those who had delivered BT in meta-regression, we did not find that superior outcomes were associated with 'higher level' qualifications. Such findings may support the assertion that BT may be suitable for shorter training and hence improve access by increasing available therapists within limited resources. We recommend further research of this question based upon our findings.

In summary, BT for depression is an effective intervention that has equal, if not better, outcomes than alternative and currently recommended therapies. Our review adds to the literature in the area as it provides a broad overview of the current evidence, reports data on recovery, dropout and explores the effect of baseline covariants in relation to depression symptom change. We recommend further research into the efficacy of behavioural treatments of depression, in particular Jacobson *et al.'s* (1996) parsimony hypothesis where the intervention is delivered by 'technicians' rather than therapists.

Declaration of Interest

None.

Note

Supplementary information accompanies this paper on the Journal's website (http://journals.cambridge.org).

References

Beck A (1976). *Cognitive Therapy and the Emotional Disorders*. International Universities Press: New York.

Beck A, Ward C, Mendelson M, Mock J, Erdbaugh J (1961). An inventory for measuring depression. *Archives of General Psychiatry* **4**, 561–571.

Begg C (1994). Publication bias. In: *The Handbook of Research Synthesis* (ed. H. Cooper and L. V. Hedges), pp. 399–409. Russell Sage Foundation: New York, NY.

Centre for Economic Performance's Mental Health Policy Group (2006). *The Depression Report. A New Deal for Depression and Anxiety Disorders*. London School of Economics: London.

Churchill R, Hunot V, Corney R, Knapp M, McGuire H, Tylee A, Wessley S (2001). A systematic review of controlled trials of the effectiveness and cost effectiveness of brief psychological treatments of depression. *Health Technology Assessment* **5** (35).

Cochrane Collaboration (2003). Review Manager (RevMan) [computer program]. Version 4.2 for Windows. The Nordic Cochrane Centre: Copenhagen.

Cole M (1983). Behavioural treatment of chronic depression. *Dissertation Abstracts International*, vol. 45, no. 3. Univ Microfilms International, US.

Cuijpers P, van Straten A, Warmerdam L (2007). Behavioural activation treatments of depression: a meta analysis. *Clinical Psychology Review* **27**, 318–326.

Cullen J, Spates R, Pagoto S, Doran N (2006). Behavioral activation treatment for major depressive disorder: a pilot investigation. *Behavior Analyst Today* **7**, 151–166.

Dimidjian S, Hollon S, Dobson K, Schmaling K, Kohlenberg R, Addis M, Gallop R, McGlinchey J, Markley D, Gollan J, Atkins D, Dunner D, Jacobson N (2006). Randomized trial of behavioural activation, cognitive therapy, and antidepressant medication in the acute treatment of adults with major depression. *Journal of Consulting and Clinical Psychology* **74**, 658–670.

Dobson K (1989). A meta analysis of the efficacy of cognitive therapy for depression. *Journal of Consulting and Clinical Psychology* **57**, 414–419.

Egger M, Davey-Smith G, Schneider M, Minder C (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal* **315**, 629–634.

Ferster C (1973). A functional analysis of depression. *American Psychologist* **28**, 857–870.

Furukawa T, Barbui C, Cipriani A, Brambilla P, Watanabe N (2006). Imputing missing standard deviations in meta-analyses can provide accurate results. *Journal of Clinical Epidemiology* **59**, 7–10.

Gallagher D, Thompson L (1982). Treatment of major depressive disorder in older adult outpatients with brief psychotherapies. *Psychotherapy* **19**, 482–490.

Gardner P, Oei T (1981). Depression and self-esteem: an investigation that used behavioural and cognitive approaches to the treatment of clinically depressed clients. *Journal of Clinical Psychology* **37**, 129–135.

Gloaguen V, Cottraux J, Cucherat M, Blackburn I (1998). A meta-analysis of the effects of cognitive therapy in depressed patients. *Journal of Affective Disorders* **49**, 59–72.

Hamilton M (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery & Psychiatry* **23**, 56–62.

Higgins J, Thompson S (2004). Controlling the risk of spurious findings from meta-regression. *Statistics in Medicine* **23**, 1663–1682.

Higgins J, Thompson S, Deeks J, Altman D (2003). Measuring inconsistency in meta-analyses. *British Medical Journal* **327**, 557–560.

Hirschfeld R, Keller M, Panico S, Arons B, Barlow D, Davidoff F, Endicott J, Froom J, Goldstein M, Gorman J, Guthrie D, Marek R, Maurer T (1997). The National Depressive and Manic-Depressive Association consensus statement on the undertreatment of

depression. *Journal of the American Medical Association* **277**, 333–340.

Hollon S, Thase M, Markowitz J (2002). Treatment and prevention of depression. *Psychological Science in the Public Interest* **3**, 1–39.

Hopko D, Lejuez C, Lepage J, Hopko S, McNeil D (2003*a*). A brief behavioral activation treatment for depression; a randomized pilot trial within an inpatient psychiatric hospital. *Behavior Modification* **27**, 458–469.

Hopko D, Lejuez C, Ruggiaro K, Eifert G (2003*b*). Contemporary behavioural activation treatments for depression: procedures, principles and progress. *Clinical Psychology Review* **23**, 699–717.

Jacobson N, Dobson K, Traux P, Addis M, Koerner K, Gollan J, Gortner E, Prince S (1996). A component analysis of cognitive-behavioural treatment of depression. *Journal of Consulting and Clinical Psychology* **64**, 295–304.

Jacobson N, Gortner E (2000). Can depression be de-medicalised in the 21st century: scientific revolutions, counter revolutions and the magnetic field of normal science. *Behaviour Research and Therapy* **38**, 103–117.

Jacobson N, Martell C, Dimijan S (2001). Behavioural activation treatment for depression: returning to contextual roots. *Clinical Psychology: Science & Practice* **8**, 255–270.

Jadad A, Moore R, Carroll D, Jenkinson C, Reynolds D, Cavaghan D, McQuay H (1996). Assessing the quality of reports of randomized clinical trials: Is blinding necessary? *Controlled Clinical Trials* **17**, 1–12.

Khan K, Kleijnen J (2002). Phase 4: Selection of Studies. Undertaking systematic reviews of research on effectiveness CRD's guidance for those carrying out or commissioning reviews. In Report 4, 2nd edn (ed. K. Khan, G. ter Riet, J. Galanville, A. Sowden and J. Kleijnen). Centre for Reviews and Dissemination, University of York, UK.

Khan K, ter Riet G, Papay J, Nixon J, Kleijnen J (2002). Phase 5: Study Quality Assesment. Undertaking systematic reviews of research on effectiveness CRD's guidance for those carrying out or commissioning reviews. In Report 4, 2nd edn (ed. K. Khan, G. ter Riet, J. Galanville, A. Sowden and J. Kleijnen). Centre for Reviews and Dissemination, University of York, UK.

Klerman G, Weissman M, Rounsaville B, Chevron E (1984). *Interpersonal Psychotherapy of Depression*. Basic Books: New York, NY.

Lewinsohn P, Graf M (1973). Pleasant activities and depression. *Journal of Consulting and Clinical Psychology* **41**, 261–295.

Lipsey M, Wilson D (1993). The efficacy of psychological, educational and behavioural treatment. *American Psychologist* **48**, 1181–1209.

Luborsky L, Mark D, Hole A (1995). Manual for supportive expressive dynamic therapy for depression: adaptation of the general SE manual. In *Psychodynamic Psychotherapy for Axis 1 Disorders* (ed. J. Barber and P. Crits-Christoph). Basic Books: New York.

Maldonado Lopez A (1982). Behavioural therapy and depression. *Revista dePsicología General y Aplicada* **37**, 31–56.

Maldonado Lopez A (1984). Behavioural therapy and depression: an experimental analysis of the interaction between cognitive and behavioural therapies and pharmacological therapy in depressed people. *Revista de Psicología General y Aplicada* **39**, 517–535.

Martell C, Addis M, Jacobson N (2001). *Depression in Context. Strategies for Guided Action*. Norton: New York.

McKendree-Smith N (1998). Cognitive and behavioural bibliotherapy for depression: an examination of efficacy and mediators and moderators of change (unpublished Ph.D. thesis). University of Alabama.

McLean P, Hakstain R (1979). Clinical depression: comparative efficacy of outpatient treatment. *Journal of Consulting and Clinical Psychology* **47**, 818–836.

McNamara K, Horan J (1986). Experimental construct validity in the evaluation of cognitive and behavioral treatments for depression. *Journal of Counselling Psychology* **33**, 23–30.

NICE (2004). *Management of Depression in Primary and Secondary Care*. National Institute for Clinical Excellence: London.

Padfield M (1976). The comparative effects of two counseling approaches on the intensity of depression among rural women of low socioeconomicstatus. *Journal of Counselling Psychology* **23**, 209–214.

Rogers C (1961). *On Becoming a Person*. Houghton Mifflin: Boston, MA.

Schulz K, Grimes D (2002). Allocation concealment in randomised trials: defending against deciphering. *Lancet* **359**, 614–618.

Scogin F, Jamison C, Gochneaur K (1989). Comparative efficacy of cognitive and behavioral bibliotherapy for mildly and moderately depressed older adults. *Journal of Consulting and Clinical Psychology* **57**, 403–407.

Singleton N, Bumpstead R, O'Brien M, Lee A, Meltzer Y (2001). *Office of National Statistics: Psychiatric Morbidity Among Adults Living in Private Households, 2000*. HMSO: London.

Skinner D (1984). Self-control of depression: A comparison of behaviour therapy and cognitive behaviour therapy. *Dissertation Abstracts International*, vol. 45, no. 3. Univ Microfilms International, US.

Stata Corporation (2003). Version 8 [program]. Stata Corporation: Texas.

Sutton A, Abrams K, Jones D (1998). Systematic reviews of trials and other studies. *Health Technology Assessment* **2** (19).

Taylor F, Marshall W (1977). Experimental analysis of a cognitive-behavioral therapy for depression. *Cognitive Therapy and Research* **1**, 59–72.

Thompson L, Gallagher D, Breckenridge J (1987). Comparative effectiveness of psychotherapies for depressed elders. *Journal of Consulting and Clinical Psychology* **55**, 385–390.

Thompson S, Higgins J (2002). How should meta-regression analyses be undertaken and interpreted? *Statistics in Medicine* **21**, 1559–1573.

WHO (2001). *The World Health Report 2001: Mental Health: New Understanding, New Hope*. World Health Organisation: Geneva.

Wilson P (1982). Combined pharmacological and behavioural treatment of depression. *Behavior Research and Therapy* **20**, 173–184.

Wilson P, Goldin J, Charbonneau P (1983). Comparative efficacy of behavioral and cognitive treatments of depression. *Cognitive Therapy and Research* **7**, 111–124.

Zeiss A, Lewinsohn P, Munoz R (1979). Nonspecific improvement effects in depression using interpersonal skills training, pleasant activity schedules or cognitive training. *Journal of Consulting and Clinical Psychology* **47**, 427–439.