Psychological Medicine, 2007, **37**, 1217–1228. © 2007 Cambridge University Press
doi:10.1017/S0033291707000062 First published online 19 February 2007 Printed in the United Kingdom

REVIEW ARTICLE

What makes self-help interventions effective in the management of depressive symptoms? Meta-analysis and meta-regression

JUDITH GELLATLY[1], PETER BOWER[2]*, SUE HENNESSY[3], DAVID RICHARDS[3], SIMON GILBODY[3] AND KARINA LOVELL[1]

[1] *Department of Nursing, Midwifery and Health Visiting, University of Manchester, UK*; [2] *National Primary Care Research and Development Centre, University of Manchester, UK*; [3] *Department of Health Sciences, University of York, UK*

ABSTRACT

Background. Although self-help interventions are effective in treating depression, less is known about the factors that determine effectiveness (i.e. moderators of effect). This study sought to determine whether the content of self-help interventions, the study populations or aspects of study design were the most important moderators.

Method. Randomized trials of the effectiveness of self-help interventions *versus* controls in the treatment of depressive symptoms were identified using previous reviews and electronic database searches. Data on moderators (i.e. patient populations, study design, intervention content) and outcomes were extracted and analysed using meta-regression.

Results. Thirty-four studies were identified with 39 comparisons. Study design factors associated with greater effectiveness were unclear allocation concealment, observer-rated outcome measures and waiting-list control groups. Greater effectiveness was also associated with recruitment in non-clinical settings, patients with existing depression (rather than those 'at risk'), contact with a therapist (i.e. guided self-help) and the use of cognitive behavioural therapy (CBT) techniques. However, only guided self-help remained significant in the multivariate analysis [regression coefficient 0·36, 95% confidence interval (CI) 0·05–0·68, $p=0·03$]. In the subset of guided studies, there were no significant associations between outcomes and the session length, content, delivery mode or therapist background.

Conclusions. The results provide some insights into moderators of self-help interventions, which might assist in the design of future interventions. However, the present study did not provide a comprehensive description, and other research methods might be required to identify factors associated with the effectiveness of self-help.

BACKGROUND

The prevalence of depression is high, and there are significant problems with accessing care (Scogin *et al*. 2003), partly due to a lack of trained therapists (Lovell & Richards, 2000). Self-help interventions can provide effective care for depression using 'health technologies' such as bibliotherapy and computer programs. Treatments without any therapist contact (so-called 'pure self-help') could have the greatest impact on access but may not be effective with depressed patients who lack motivation and confidence. In the UK, the term 'guided self-help' is used to refer to self-help interventions with minimal therapist contact that might

* Address for correspondence: Dr Peter Bower, National Primary Care Research and Development Centre, University of Manchester, Manchester M13 9PL, UK.
(Email: peter.bower@manchester.ac.uk)

provide the optimal balance between efficiency and effectiveness.

There is developing evidence that self-help interventions can be effective (McKendree-Smith *et al.* 2003; Anderson *et al.* 2005), and they have been proposed as a key part of stepped care models of depression care in both the USA and the UK (Scogin *et al.* 2003; National Collaborating Centre for Mental Health, 2005). However, evidence of effectiveness is not uniformly positive (Richards *et al.* 2003; Proudfoot *et al.* 2004; Mead *et al.* 2005). Significant variation in clinical effectiveness beyond that expected by chance is known as heterogeneity. Heterogeneity may be reliably associated with factors such as the quality of study design, patient populations and the context of care, or may relate to differences in the design of self-help interventions themselves. Such factors are known as moderators of treatment effect (Baron & Kenny, 1986). Understanding the contribution of different moderators is crucial for the interpretation of study results and the design of future interventions. Meta-regression is a technique that can be used to examine moderators. Whereas a meta-analysis combines results in order to obtain a single summary effect size, meta-regression aims to relate the size of effect to one or more characteristics of the studies involved (Thompson & Higgins, 2002).

Several recent reviews have been conducted in this area (Bower *et al.* 2001; McKendree-Smith *et al.* 2003; Den Boer *et al.* 2004; Anderson *et al.* 2005). Because of the paucity of studies, most have been unable to quantify key moderators. A recent review including 14 studies suggested that self-report assessments were associated with lower effect size but that severity of illness did not influence the results (Den Boer *et al.* 2004). However, there were insufficient studies available for a more comprehensive analysis.

The policy focus on self-help interventions means that additional studies are published frequently, and previous reviews must be updated. The aim of the current study was to provide a systematic review of the randomized controlled trial literature to determine intervention, population and study design factors that moderate the treatment effect of self-help interventions for depression.

METHOD

Objective

To use meta-regression to determine intervention, population and study design factors that moderate the treatment effect of self-help interventions for depression.

Inclusion and exclusion criteria

Because a key focus of the study was to examine the relationship between variation in study characteristics and outcomes, the inclusion criteria were deliberately kept relatively broad (Gotzsche, 2000).

Study design

Only randomized controlled trials were eligible.

Population

Depressive disorders, depressive symptoms or psychosocial problems (e.g. stress), where depressive symptoms were measured as a primary outcome, were all included, as were studies of the prevention of depression in patients at risk. All settings were also eligible, including community, primary care, specialist out-patient and in-patient and non-clinical settings.

Interventions

There is no agreed definition of self-help, or clarity about the demarcation between self-help and conventional brief therapy. In general, the interventions of interest were designed to assist patients in the treatment of their depressive symptoms, using a health technology such as written information, audiotape, videotape or computer presentation. Interventions were designed to be conducted predominantly independent of professional or paraprofessional contact. There is no consensus concerning the appropriate amount of therapist contact for a treatment to be described as 'self-help'. In the UK, current treatment length for conventional brief therapy is 6–8 sessions (National Collaborating Centre for Mental Health, 2005), and for the present review the maximum amount of professional input for self-help interventions was defined as half that found at the lower range for conventional therapy (i.e. 3 hours or less). This definition would include studies meeting published definitions of 'self-administered

therapy', 'predominantly self-help' and 'minimal contact therapy' (Newman *et al*. 2003).

Literature search strategies

There were several methods of identifying studies for the review. A number of previous reviews had been conducted in this area (Scogin *et al*. 1990; Gould and Clum, 1993; Marrs, 1995; Cuijpers, 1997; Bower *et al*. 2001; McKendree-Smith *et al*. 2003; Den Boer *et al*. 2004; Anderson *et al*. 2005) and these were initially examined for relevant studies. A comprehensive search (without language restriction) of the Cochrane Collaboration's register of controlled trials (CENTRAL) was then undertaken to identify any studies missed by these previous reviews. The search of CENTRAL was then augmented by searches of Medline, EMBASE, CINAHL and PsycINFO for the years 2002–2005 to identify any recent additional literature that might not have been identified within CENTRAL. Searches used a mix of subject headings and free text terms for self-help interventions and depression, combined with a conventional randomized controlled trials search filter (an example search is available in the online Appendix).

Study selection and data extraction

Titles and abstracts were judged by individual researchers who identified potentially eligible studies. Full text articles were then discussed by the research team and assessed against the study inclusion criteria. When studies were judged as eligible, data were extracted from published reports by two researchers working independently, with disagreements resolved by discussion or contact with authors. For the purposes of this review, data extraction was restricted to comparisons of self-help interventions *versus* 'control' groups, and comparisons with other treatments (e.g. conventional therapy) were excluded as they were comparatively rare. Control groups included a range of comparators such as attention placebos, waiting list and usual care. Data on the codes are presented in Table 1.

Study design moderators

Three measures of internal validity were used. The first measure was concealment of allocation, which was a judgement of whether random allocation was conducted such that judgements of eligibility for the trial were made independent of knowledge of the allocation sequence (Schulz & Grimes, 2002), a known moderator of effect size (Schulz *et al*. 1995).

The other measures of internal validity were the use of an intention-to-treat analysis (Hollis & Campbell, 1999) and levels of attrition (Cook & Campbell, 1979; Smith *et al*. 1980), using an arbitrary threshold of 80% follow-up. Judgements were made on the basis of published information; where information was missing, studies were rated as 'low' quality for that category.

Two other study design issues were coded. Given that the type of assessment (self-report *versus* external assessor) was found to be a moderator of outcome in a previous review (Den Boer *et al*. 2004), this was also coded. Finally, control groups were coded as 'attention placebos', 'waiting list' or 'usual care'.

Study context and population moderators

Recruitment context was coded as either 'non-clinical' (mainly media advertisements or unselected surveys in the community) or 'clinical' (e.g. referral from a professional in a primary care or mental health setting) or mixed. Patient populations were coded as either patients 'at risk' for depression (e.g. recruited from particular groups, such as pregnant women, based on likely risk) or patients recruited on the basis of existing depressive symptoms. Studies including patients 'at risk' did so on the assumption that a proportion of 'at risk' patients would develop more severe symptoms during the study, and that the proportion would be reduced in those receiving self-help. Patients were also coded as to whether they had received a psychiatric diagnosis of depression according to a standardized scheme such as DSM-IV or ICD-10.

Intervention moderators

The type of health technology was coded as either 'written bibliotherapy' or 'computer technology'. The theoretical model underlying the intervention was coded as either 'educational' (when the intervention was restricted to information provision) or 'cognitive behavioural therapy' (CBT) (when the intervention included specific descriptions of CBT

Table 1. *Descriptive data on moderator variables*

Moderator	Subcategories	*n* (%)
All studies		
Recruitment setting	Non-clinical settings	22 (56)
	Clinical settings and mixed	17 (44)
Population type	Patients at risk of depression	9 (23)
	Patients with current symptoms of depression	30 (77)
Diagnosis	Patients without standardized psychiatric diagnosis	29 (74)
	Patients with standardized psychiatric diagnosis	10 (26)
Concealment of allocation	Adequate	14 (36)
	Inadequate or not clear	25 (64)
Intention-to-treat analysis	Yes	24 (62)
	No	15 (38)
Attrition	$\leqslant$20%	17 (44)
	>20%	22 (56)
Outcome assessment	Self-report	35 (90)
	External assessment	4 (10)
Control group	Waiting list	23 (59)
	Usual care	11 (28)
	Attention placebo	5 (13)
Technology	Written materials	29 (74)
	Computer materials	10 (26)
Theoretical model	Educational	8 (21)
	CBT	31 (79)
Guidance	'Pure' self-administered interventions	15 (39)
	'Guided' interventions	24 (61)
Guided interventions		
Number of sessions	As a quantitative variable (mean, S.D.)	2·8 (2·8)
Personnel	Paraprofessional worker	21 (88)
	Professional workers	3 (12)
Content of the guidance	Monitoring	17 (71)
	Supportive	7 (29)
Mode of guidance	Face to face	5 (21)
	Other	19 (79)

CBT, Cognitive behavioural therapy; S.D., standard deviation.

techniques). Studies were coded as 'pure' or 'guided' self-help based on the presence or absence of therapist contact as part of the intervention. In the subset of 'guided' studies, the background of the therapist was coded as either 'professional' (postgraduate mental health qualification) or 'paraprofessional' (no postgraduate mental health qualification). The mode of delivery of guidance was coded as 'face-to-face' or 'other' (which included telephone, email and written contact). The content of the guidance was coded as 'monitoring', where the focus was on checking that patients had used the materials, assisting with queries about use, or where there was a specific statement that therapeutic techniques such as counselling were not used. Interventions were coded as having 'supportive' content when the intervention involved more than simple monitoring, which might include advice from the therapist about specific problems, motivation or support.

Outcomes

Outcome measures in the studies included self-report and external assessment, and ranged from measures of depressive symptoms to related psychological processes such as cognitive styles. Continuous measures were translated to a standardized effect size, that is mean of intervention group minus mean of control group, divided by the pooled standard deviation. Outcomes reported as dichotomous variables were translated to a standardized effect size using the logit transformation (Lipsey & Wilson, 2001).

Many studies included multiple outcome measures, and used multiple follow-up points. The primary analysis was restricted to measures of depressive symptoms or diagnoses, and measures of related psychological processes such as cognitive styles were excluded. However, some studies still reported multiple relevant measures, and their inclusion would violate

statistical assumptions about independence. In these cases, it is possible to average over multiple dependent measures, or use a single measure chosen *a priori* (Lipsey & Wilson, 2001). The latter approach was used in the current study. The frequency of use of depression measures in the entire sample was calculated, and the most frequently used measure available within any study was chosen for the primary analysis (which usually involved validated measures such as the Beck Depression Inventory, the Hamilton Rating Scale for Depression, the Centre for Epidemiological Studies – Depression Scale and the Hospital Anxiety and Depression Scale). To assess the possible effect of this choice, a secondary sensitivity analysis used those outcome measures within each trial that gave the highest and lowest standardized effect size.

Data extraction was restricted to short-term outcomes, as long-term outcomes were infrequently reported. Where multiple short-term follow-ups were reported, the analysis used the longest recorded follow-up within a 6-month period. This provided a reasonable degree of consistency in terms of the length of follow-up across studies.

Analysis

Meta-regression extends conventional meta-analysis to estimate the extent to which one or more moderators in each study explain heterogeneity in treatment effect. Analyses were conducted in Stata version 9 (Stata Corporation, College Station, TX, USA), using the METAN and METAREG macros. The primary analyses were conducted on all studies initially, and then restricted to the subset of 'guided' studies when examining the moderating effects of different types of guidance.

Where studies reported two parallel comparisons with a control group (e.g. two types of bibliotherapy *versus* waiting list), both comparisons were included separately, but sample sizes in the control group were halved to ensure that they were not double counted in the meta-analysis. One study reported analyses of a trial split by site (Rahe *et al.* 2002), and these were treated as two separate comparisons. Two studies used factorial designs (Reid *et al.* 2002; Little *et al.* 2004), although in both cases only one of the factorial comparisons was related to the study question and included in the analyses.

The initial meta-analyses used random effects modelling (Sutton *et al.* 1998) to provide a conventional pooled effect size. Random effects modelling is based on the assumption that estimates of effects are distributed randomly. Such models are superior in cases where studies are markedly heterogeneous, and are thus appropriate in the present context. Heterogeneity was measured using Cochran's test of heterogeneity, but this test provides only an assessment of the significance of heterogeneity, not its magnitude, is susceptible to the number of trials in the meta-analysis, and is difficult to interpret when comparing meta-analyses. Therefore, the I^2 statistic, which estimates the percentage of total variation across studies that can be attributed to heterogeneity rather than chance, was also calculated. As a guide, I^2 values of 25% may be considered 'low', 50% 'moderate' and 75% 'high' (Higgins *et al.* 2003).

The main analysis of moderators of treatment effect used random effects meta-regression, which provided estimates of the relationships between moderators and outcomes in the form of a regression coefficient and related confidence intervals (CIs). This regression coefficient represents the difference in effect between trials rated on the two levels of categorical moderator variables (such as 'pure' and 'guided' self-help) or the change in effect associated with a one-unit increase in a continuous moderator variable (such as average session length within a trial). The permutation test was used to calculate p values (using 1000 Monte-Carlo simulations) to reduce the chance of spurious false-positive findings (Higgins & Thompson, 2004). Initial univariate analyses were followed by estimation of a multivariate model. The amount of heterogeneity explained by moderator variables was examined by reductions in the I^2 statistic when the moderator was entered into the analysis.

Publication bias refers to the situation when published research is systematically unrepresentative of the population of completed studies (Rosenthal, 1979), and was assessed through a funnel plot (Egger *et al.* 1997), which plotted effect size against the standard error of the effect size. When an absence of small studies with small effect sizes causes asymmetry in the plot, publication bias may be present.

RESULTS

We identified 34 published studies reporting 39 relevant comparisons. Fourteen studies were identified from previous reviews, and 20 were located from the 22 383 studies identified by the new electronic database searches (the flowchart is available in the online Appendix).

Meta-analysis

Overall, the effect of self-help interventions was 'medium' according to current convention, with a pooled standardized mean difference of 0·43 and 95% CI 0·30–0·57 (Cohen, 1988; Lipsey, 1990). The results are plotted in Fig. 1. The variation in effect size (I^2) attributable to heterogeneity was 77·3%. When the analysis was restricted to those studies using 'guided self-help', the pooled standardized mean difference was 'large' according to current convention (0·80, 95% CI 0·58–1·01), and the variation in effect size attributable to heterogeneity was 68·3%.

Meta-regression

The results of the analyses of all studies are shown in Table 2.

Study context and population moderators

Studies involving patients recruited in non-clinical settings reported a significantly higher effect size than those in clinical populations. Two comparisons involved populations with 'stress', and nine were with patients 'at risk' for depression rather than suffering from existing problems. Studies with populations 'at risk' for depression showed a significantly smaller effect size than those in patients with existing problems.

Study design moderators

There was no association between effect size and intention-to-treat analyses, or attrition rates. However, effects were significantly lower when allocation was adequately concealed, when assessment was through self-report, and when the control group was 'usual care' or 'attention placebo' as opposed to 'waiting list'.

Intervention moderators

Studies using a 'guided' model reported a significantly higher effect size than those using 'pure' self-help. Interventions based on a CBT model reported a significantly higher effect size than those based on an educational model. There was no significant difference between the effect of self-help based on written bibliotherapy and those based on information technology.

When the analysis was restricted to 'guided' studies (Table 3), there were no significant relationships between effect size and personnel type, content of the guidance, mode of the guidance or number of sessions.

Multivariate analyses

A multivariate analysis was also conducted using any moderators with associations significant in the univariate analyses across all studies (i.e. concealment of allocation, type of assessment, type of control group, recruitment setting, population type, theoretical model, and 'guided' *versus* 'pure' self-help). Only 'guided' *versus* 'pure' remained significant in the multivariate analysis (regression coefficient 0·36, 95% CI 0·05–0·68, $p=0·03$). No multivariate analysis was conducted restricted to 'guided' studies as there were no significant moderators in the univariate analyses.

Secondary analyses

As noted earlier, most studies reported multiple outcomes, and the most commonly used outcomes were used in the primary analysis. An *a priori* secondary analysis examined the effect of including those outcome measures that gave the highest and lowest estimates of effect within each study, as opposed to the most commonly reported measure. When the highest effect sizes were used, there were no changes in the statistical significance of associations reported in Tables 2 and 3. When the lowest effect sizes were used, recruitment context, theoretical model, concealment of allocation and type of outcome assessment were no longer significant moderators, whereas diagnosis became significant.

A *post hoc* secondary analysis was also conducted to assess the effect of one study that was based on an adolescent population, and demonstrated a highly positive effect size (Ackerson *et al.* 1998). The statistical significance of the results reported in Tables 2 and 3 were not influenced by the exclusion of this study.

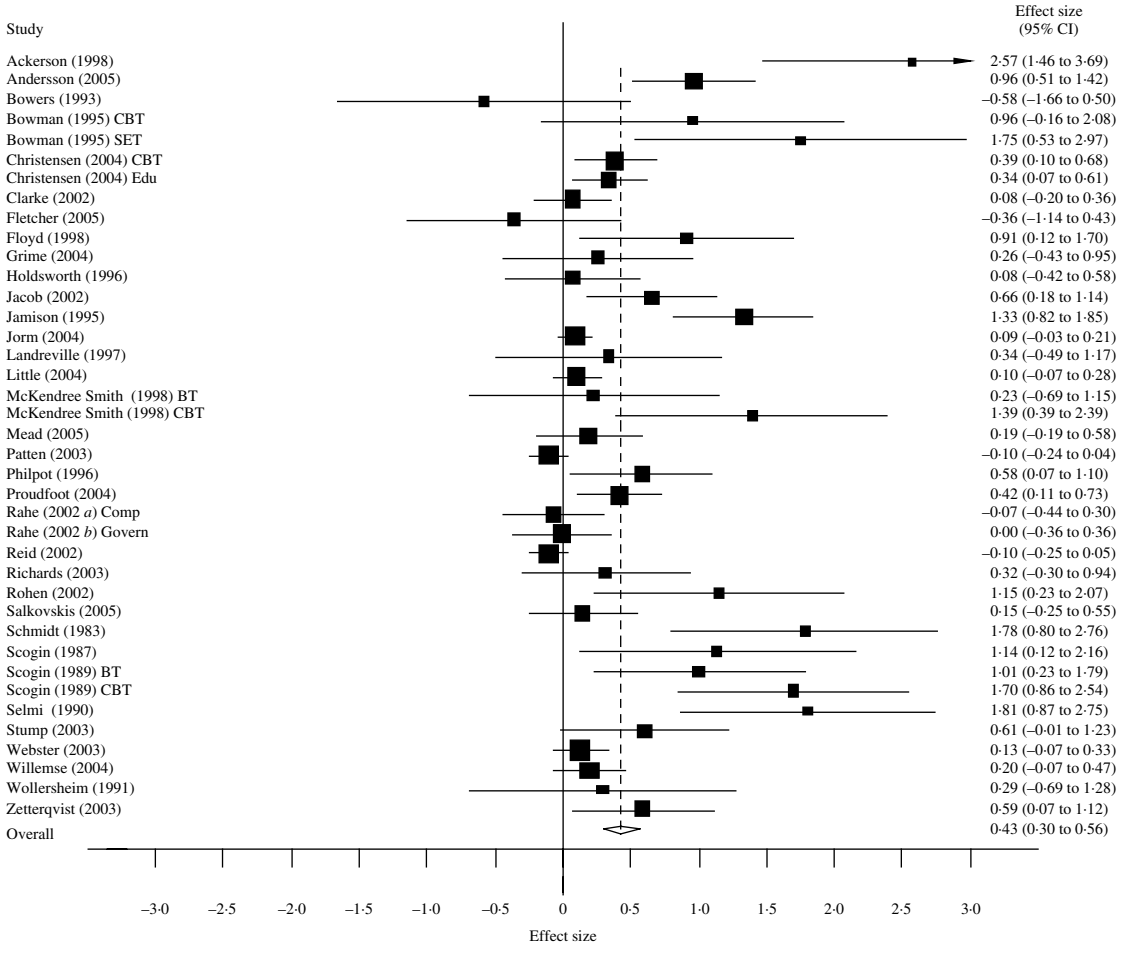

FIG. 1. Random-effects meta-analysis of estimates of effect size. CBT, Cognitive behavioural therapy; SET, self-examination therapy; BT, behavioural therapy; Edu, educational website; Comp, computer industries sample; Govern, city government sample.

Analysis of publication bias

The funnel plot showed significant asymmetry, which suggested that publication bias might be present (Fig. 2). The test for asymmetry among all studies was positive using Egger's test ($p < 0{\cdot}001$). However, publication bias was only evident among 'guided' self-help studies ($p < 0{\cdot}0001$), but was non-significant in 'pure' self-help studies ($p = 0{\cdot}39$).

DISCUSSION

The results of the overall meta-analysis concerning the effectiveness of self-help interventions are lower than previous studies, which have reported effect sizes of 1·36 (Anderson *et al*. 2005) and 0·84 (Den Boer *et al*. 2004). The current review included significantly more studies than previous reviews, partly because new studies were published in the time between reviews, and partly because the current review used more liberal inclusion criteria. The study took a liberal approach to inclusion because the focus was on moderators of treatment effects. This led to the inclusion of studies with a wider range of interventions, including simple provision of educational booklets as well as more complex self-help technologies. A wider variety of patients were also included, such as those with diagnosed depression, depressive

Table 2. *Univariate analyses of moderators of self-help interventions*

Variable	Category 1	Category 2	Regression coefficient (95% CI)	*p* value[a]	I^2 inconsistency statistic[b] Q statistic
Concealment of allocation	Inadequate or unclear ES = 0·72	Adequate ES = 0·21	−0·41 (−0·76 to −0·06)	0·01	77·4% 163·5 (37 df), $p = 0·000$
Intention to treat analysis	No ES = 0·58	Yes ES = 0·39	−0·08 (−0·48 to 0·33)	0·67	77·6% 165·3 (37 df), $p = 0·000$
Attrition	>20% ES = 0·36	⩽20% ES = 0·52	0·12 (−0·25 to 0·49)	0·46	77·5% 164·5 (37 df), $p = 0·000$
Outcome assessment	Self-report ES = 0·35	External assessment ES = 1·47	1·08 (0·40 to1·77)	0·00	72·7% 135·4 (37 df), $p = 0·000$
Type of control group[c]	Waiting list ES = 0·82	Usual care/attention placebo ES = 0·14	−0·56 (−0·87 to −0·25)	0·00	68·9% 119·0 (37 df), $p = 0·000$
Recruitment setting	Non-clinical setting ES = 0·66	Clinical and mixed setting ES = 0·22	−0·37 (−0·73 to −0·02)	0·03	77·5% 164·1 (37 df), $p = 0·000$
Population type	At risk of depression ES = 0·07	Existing depression ES = 0·62	0·46 (0·11 to 0·82)	0·00	72·6% 135·1 (37 df), $p = 0·000$
Diagnosis	No standardized diagnosis ES = 0·33	Standardized diagnosis ES = 0·70	0·35 (−0·09 to 0·78)	0·07	73·3% 138·7 (37 df), $p = 0·000$
Technology	Written ES = 0·46	Computer ES = 0·38	−0·09 (−0·50 to 0·32)	0·63	77·9% 167·3 (37 df), $p = 0·000$
Theoretical model	Educational ES = 0·11	CBT ES = 0·61	0·43 (0·05 to 0·80)	0·02	75·8% 152·7 (37 df), $p = 0·000$
'Guided' or 'pure' self-help	'Pure' self-help ES = 0·06	'Guided' self-help ES = 0·80	0·63 (0·36 to 0·90)	0·00	61·3% 95·7 (37 df), $p = 0·000$

CI, Confidence interval; ES, effect size; df, degrees of freedom; CBT, cognitive behavioural therapy.

[a] Based on 1000 Monte Carlo simulations.

[b] I^2 statistic describes the percentage of total variation across studies that is due to heterogeneity rather than chance. A value of 0% indicates no observed heterogeneity, and larger values show increasing heterogeneity. The I^2 value in the table reflects the amount of heterogeneity remaining after accounting for each moderator in the row of the table.

[c] Usual care and attention placebo collapsed into a single category as the estimates of effects were similar (−0·15 and −0·13).

Table 3. *Univariate analyses of moderators of guided self-help interventions*

Variable	Category 1	Category 2	Regression coefficient (95% CI)	*p* value[a]	I^2 inconsistency statistic[b] Q statistic
Number of sessions[c]	Continuous variable		−0·12 (−0·29 to 0·04)	0·15	67·0% 57·7 (19 df, $p < 0·000$)
Personnel	Paraprofessional ES = 0·86	Professional ES = 0·50	−0·36 (−1·07 to 0·35)	0·31	68·4% 69·6 (22 df, $p < 0·000$)
Content of the guidance	Monitoring ES = 0·78	Supportive ES = 0·83	0·03 (−0·52 to 0·58)	0·92	68·3% 69·3 (22 df, $p < 0·000$)
Mode of guidance	Other ES = 0·89	Face to face ES = 0·49	−0·38 (−0·97 to 0·22)	0·20	68·4% 69·7 (22 df, $p < 0·000$)

CI, Confidence interval; ES, effect size; df, degrees of freedom.

[a] Based on 1000 Monte Carlo simulations.

[b] I^2 statistic describes the percentage of total variation across studies that is due to heterogeneity rather than chance. A value of 0% indicates no observed heterogeneity, and larger values show increasing heterogeneity. The I^2 value in the table reflects the amount of heterogeneity remaining after accounting for each moderator in the row of the table.

[c] Based on $n = 21$.

symptoms, stress, and those 'at risk'. The difference in the overall effect size between the current review and previous studies may reflect variation in inclusion criteria, and these issues must be kept in mind when interpreting the results.

However, the overall estimate of effect was of less interest than the analysis of moderating

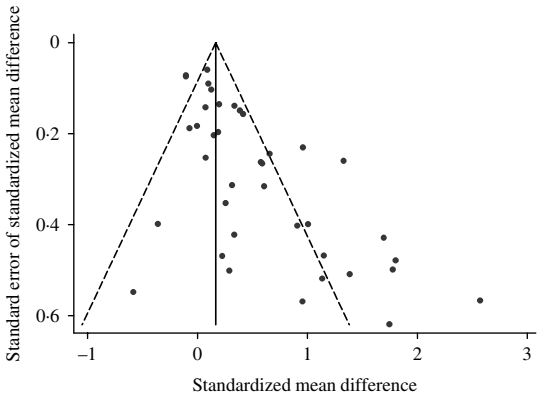

FIG. 2. Funnel plot of all studies. Graph of asymmetrical funnel plot suggestive of publication bias. Egger's regression test (zero if unbiased) = 2·41 (95% CI 1·60 to 3·22), $p < 0·05$.

variables, and the present study extends the results of previous reviews by giving a more comprehensive examination of moderators. It should be noted that only guided self-help was a significant moderator of outcomes in the multivariate analysis. The interpretation of multivariate regression is complex, as the analysis takes account of the relationships between moderator variables, as well as those between the moderators and the outcome. If a variable predicts outcome in a univariate analysis but fails to reach significance in a multivariate analysis, this may not mean that it is unimportant. Nevertheless, the following discussion of variables that are only significant in the univariate analyses should be viewed with some caution, as their statistical significance may reflect statistical confounding between moderators.

Concealment of allocation was a key moderator of treatment effect, replicating earlier work (Schulz *et al*. 1995). Similarly, the effect of type of outcome assessment supported a previous review of self-help (Den Boer *et al*. 2004). The mechanism underlying the moderating effect of the type of control group is unclear. Few studies provided detailed information about the services available to those in the control groups. However, the size of the effect does suggest caution in generalizing the results from waiting-list studies.

The finding that the interventions had a smaller effect in patients recruited from clinical settings rather than volunteers contradicts a previous review (Den Boer *et al*. 2004) but is in accord with a recent systematic review of brief therapy for depression (Churchill *et al*. 2002). It is noteworthy that the evidence base used for self-administered treatments within the National Institute for Clinical Excellence (NICE) depression guidelines was almost entirely conducted in volunteer populations.

Overall, self-help interventions may be more effective with existing problems rather than in a preventive capacity (the difference in effect between patients with and without a diagnosed depressive disorder was marked, although statistically non-significant). It is generally supposed that self-help interventions are of greatest benefit in milder problems, but clearly the capacity to benefit in populations at risk may be smaller, such that it may be difficult to show additional benefit from *any* intervention compared to no treatment.

The univariate results suggested that self-help interventions should be based on CBT principles rather than education. This may be unsurprising, but has important implications, as simpler educational materials may have greater applicability to some populations, such as those with poor health literacy. It is still unclear whether other theoretical models would show equal efficacy to CBT, and it is unknown what the optimal mix is between cognitive and behavioural components.

The multivariate analysis gave clear evidence of the superiority of 'guided' interventions, which supports current treatment guidelines (National Collaborating Centre for Mental Health, 2005). However, when analyses were restricted to 'guided' studies, there was no clear relationship between number of sessions and outcome, which means that no specific recommendations on the type of guidance can be given. It should be noted that the range in session length is necessarily restricted when considering self-administered interventions, and power to detect associations was limited.

There was no clear advantage associated with contact that involved 'supportive' guidance beyond 'monitoring'. However, this finding should be treated with caution. The coding of this attribute was problematic because the exact nature of the contact was not always made clear within published reports.

The study only tested a small number of moderators that were felt to be likely sources of heterogeneity, as there were concerns that testing large numbers of moderators would inflate the chance of false-positive findings. However, it is possible that other variables might be important. It should also be noted that the reliability of the coding used in the study was not assessed quantitatively.

Limitations of the study

It is possible that additional studies would have been identified if the electronic database search had been extended to cover databases from inception, as previous reviews might have used different inclusion criteria or missed relevant studies. However, the search did use the CENTRAL database in the Cochrane Library, which is a comprehensive database of randomized controlled trials identified from a large number of databases without time restriction, as well as the results of handsearches of journals conducted by the Cochrane collaboration (Dickersin *et al.* 2002). It has been suggested that traditional exhaustive search strategies using multiple databases are not cost-effective, as this has already been done for CENTRAL (Royle & Waugh, 2005). It is possible that the study failed to examine sufficient grey literature sources and other methods for the identification of unpublished studies (such as contact with authors) and the funnel plot did suggest that publication bias might be present. However, there are a number of interpretations of the asymmetrical plot other than publication bias (Shadish & Baldwin, 2005).

The analysis is limited because it is restricted to short-term outcomes. Moderators of short-term benefits may differ from those that predict enduring effects. However, studies using waiting-list controls did not generally report long-term outcomes with appropriate control groups.

The technique of meta-regression has several limitations (Thompson & Higgins, 2002). The analysis only represents an observational association because meta-regression across trials does not have the benefits of randomization. Primary randomized controlled trials and 'dismantling' studies (Jacobson *et al.* 1996) are optimal methods for examining moderators, but may not always be feasible.

Power to detect associations is limited by the number of available studies (Lambert *et al.* 2002), and limitations of power are especially important in the analysis of studies using the 'guided' model ($n=24$). In a small sample, outliers might have a large influence, and multivariate models might be highly sensitive to the inclusion of particular variables. Finally, some moderators (such as the number of guidance sessions) might demonstrate a relationship within trials that is not present in differences between trials (Thompson & Higgins, 2002).

Implications of the study

The results would support the NICE recommendation about the appropriate treatment being materials based on CBT with guidance from a health professional, although the length or nature of that guidance is less clear. Other associations should be interpreted with caution, but might suggest that the particular technology used is less crucial, and that a range of delivery methods, professionals and content can be used successfully. The current results also identify key population and methodological moderators. These may be useful in determining whether the effects of self-help shown in a study completed in one context can be reasonably generalized to another.

Although the meta-regression provides some useful insights, the multivariate analysis only identified a single significant moderator. A key issue for the future is the comparison of different methods of investigation in the analysis of moderators of treatment effect. Qualitative research (Elliott *et al.* 1999) and process methods (Elliott, 1984) may be required to provide a more fine-grained analysis and provide more detail concerning the optimal methods of delivering self-help interventions.

ACKNOWLEDGEMENTS

This study was funded by the Medical Research Council Brain Sciences Initiative.

DECLARATION OF INTEREST

None.

NOTE

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

REFERENCES

Ackerson, J., Scogin, F., McKendree-Smith, N. & Lyman, R. (1998). Cognitive bibliotherapy for mild and moderate adolescent depressive symptomatology. *Journal of Consulting and Clinical Psychology* **66**, 685–690.

Anderson, L., Lewis, G., Araya, R., Elgie, R., Harrison, G., Proudfoot, J., Schmidt, U., Sharp, D., Weightman, A. & Williams, C. (2005). Self-help books for depression: how can practitioners and patients make the right choice? *British Journal of General Practice* **55**, 387–392.

Baron, R. & Kenny, D. (1986). The moderator–mediator distinction in social psychological research: conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology* **51**, 1173–1182.

Bower, P., Richards, D. & Lovell, K. (2001). The clinical and cost-effectiveness of self-help treatments for anxiety and depressive disorders in primary care: a systematic review. *British Journal of General Practice* **51**, 838–845.

Churchill, R., Hunot, V., Corney, R., Knapp, M., McGuire, H., Tylee, A. & Wessely, S. (2002). A systematic review of controlled trials of the effectiveness and cost-effectiveness of brief psychological treatments for depression. *Health Technology Assessment* Vol. 5, No. 35.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioural Sciences* (2nd edn). Lawrence Erlbaum: New Jersey.

Cook, T. & Campbell, D. (1979). *Quasi-Experimentation – Design and Analysis Issues for Field Settings*. Rand McNally: Chicago.

Cuijpers, P. (1997). Bibliotherapy in unipolar depression: a meta-analysis. *Journal of Behaviour Therapy and Experimental Psychiatry* **28**, 139–147.

Den Boer, P., Wiersma, D. & Van Den Bosch, R. (2004). Why is self-help neglected in the treatment of emotional disorders? A meta-analysis. *Psychological Medicine* **34**, 959–971.

Dickersin, K., Manheimer, E., Wieland, S., Robinson, K., Lefebvre, C., McDonald, S. & CENTRAL Development Group (2002). Development of the Cochrane Collaboration's CENTRAL Register of Controlled Clinical Trials. *Evaluation and the Health Professions* **25**, 38–64.

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.

Elliott, R. (1984). A discovery-oriented approach to significant change events in psychological therapies: interpersonal process recall and comprehensive process analysis. In *Patterns of Change: Intensive Analysis of Psychological Therapies Process* (ed. L. Rice & L. Greenberg), pp. 249–286. Guilford Press: London.

Elliott, R., Fischer, C. & Rennie, D. (1999). Evolving guidelines for publication of qualitative research studies in psychology and related fields. *British Journal of Clinical Psychology* **38**, 213–229.

Gotzsche, P. (2000). Why we need a broad perspective on meta-analysis. *British Medical Journal* **321**, 585–586.

Gould, R. & Clum, G. (1993). A meta-analysis of self-help treatment approaches. *Clinical Psychology Review* **13**, 169–186.

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.

Hollis, S. & Campbell, F. (1999). What is meant by intention to treat analysis? Survey of published randomised controlled trials. *British Medical Journal* **319**, 670–674.

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

Lambert, P., Sutton, A., Abrams, K. & Jones, D. (2002). A comparison of summary patient-level covariates in meta-regression with individual patient data meta-analysis. *Journal of Clinical Epidemiology* **55**, 86–94.

Lipsey, M. (1990). *Design Sensitivity: Statistical Power for Experimental Research*. Sage: Newbury Park.

Lipsey, M. & Wilson, D. (2001). *Practical Meta-Analysis*. Sage: Newbury Park.

Little, P., Dorward, M., Warner, G., Moore, M., Stephens, K., Senior, J. & Kendrick, T. (2004). Randomised controlled trial of effect of leaflets to empower patients in consultations in primary care. *British Medical Journal* **328**, 441–444.

Lovell, K. & Richards, D. (2000). Multiple Access Points and Levels of Entry (MAPLE): ensuring choice, accessibility and equity for CBT services. *Behavioral and Cognitive Psychotherapy* **28**, 379–391.

Marrs, R. (1995). A meta-analysis of bibliotherapy studies. *American Journal of Community Psychology* **23**, 843–870.

McKendree-Smith, N., Floyd, M. & Scogin, F. (2003). Self-administered treatments for depression: a review. *Journal of Clinical Psychology* **59**, 275–288.

Mead, N., MacDonald, W., Bower, P., Lovell, K., Richards, D. & Bucknall, A. (2005). The clinical effectiveness of guided self-help versus waiting list control in the management of anxiety and depression: a randomised controlled trial. *Psychological Medicine* **35**, 1633–1643.

National Collaborating Centre for Mental Health (2005). *Depression: Management of Depression in Primary and Secondary Care – NICE Guidance*. National Institute for Clinical Excellence (www.nice.org.uk/page.aspx?o = 235213). Accessed 19 January 2005.

Newman, M., Erickson, T., Przeworski, A. & Dzus, E. (2003). Self-help and minimal-contact therapies for anxiety disorders: is human contact necessary for therapeutic efficacy? *Journal of Clinical Psychology* **59**, 251–274.

Proudfoot, J., Ryden, C., Everitt, B., Shapiro, D., Goldberg, D., Mann, A., Tylee, A., Marks, I. & Gray, J. (2004). Clinical efficacy of computerised cognitive-behavioural therapy for anxiety and depression in primary care: randomised controlled trial. *British Journal of Psychiatry* **185**, 46–54.

Rahe, R., Taylor, C., Tolles, R., Newhall, M., Veach, T. & Bryson, S. (2002). A novel stress and coping workplace program reduces illness and healthcare utilization. *Psychosomatic Medicine* **64**, 278–286.

Reid, M., Glazener, C., Murray, G. & Taylor, G. (2002). A two-centred pragmatic randomised controlled trial of two interventions of postnatal support. *British Journal of Obstetrics and Gynaecology* **109**, 1164–1170.

Richards, A., Barkham, M., Cahill, J., Richards, D., Williams, C. & Heywood, P. (2003). PHASE: a randomised, controlled trial of supervised self-help cognitive behavioural therapy in primary care. *British Journal of General Practice* **53**, 764–770.

Rosenthal, R. (1979). The 'File Drawer Problem' and tolerance for null results. *Psychological Bulletin* **86**, 638–641.

Royle, P. & Waugh, N. (2005). A simplified search strategy for identifying randomised controlled trials for systematic reviews of health care interventions: a comparison with more exhaustive strategies. *BMC Medical Research Methodology* **5**, 23.

Schulz, K., Chalmers, I., Hayes, R. & Altman, D. (1995). Empirical evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *Journal of the American Medical Association* **273**, 408–412.

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

Scogin, F., Bynum, J., Stephens, G. & Calhoon, S. (1990). Efficacy of self-administered treatment programs: meta-analytic review. *Professional Psychology: Research and Practice* **21**, 42–47.

Scogin, F., Hanson, A. & Welsh, D. (2003). Self-administered treatment in stepped-care models of depression treatment. *Journal of Clinical Psychology* **59**, 341–349.

Shadish, W. & Baldwin, S. (2005). Effects of behavioral marital therapy: a meta-analysis of randomized controlled trials. *Journal of Consulting and Clinical Psychology* **73**, 6–14.

Smith, M., Glass, G. & Miller, T. (1980). *The Benefits of Psychotherapy*. Johns Hopkins University Press: Baltimore.

Sutton, A., Abrams, K., Jones, D., Sheldon, T. & Song, F. (1998). Systematic reviews of trials and other studies. *Health Technology Assessment* Vol. 2, No. 19.

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