

# Psychological treatment of depression in primary care: a meta-analysis

Pim Cuijpers, Annemieke van Straten, Anneke van Schaik and Gerhard Andersson

## ABSTRACT

### Background

Although most depressive disorders are treated in primary care and several studies have examined the effects of psychological treatment in primary care, hardly any meta-analytic research has been conducted in which the results of these studies are integrated.

### Aim

To integrate the results of randomised controlled trials of psychological treatment of depression in adults in primary care, and to compare these results to psychological treatments in other settings.

### Design of study

A meta-analysis of studies examining the effects of psychological treatments of adult depression in primary care.

### Setting

Primary care.

### Method

An existing database of studies on psychological treatments of adult depression that was built on systematic searches in PubMed, PsychINFO, EMBASE, and Dissertation Abstracts International was used. Randomised trials were included in which the effects of psychological treatments on adult primary care patients with depression were compared to a control condition.

### Results

In the 15 included studies, the standardised mean effect size of psychological treatment versus control groups was 0.31 (95% CI = 0.17 to 0.45), which corresponds with a numbers-needed-to-treat (NNT) of 5.75. Studies in which patients were referred by their GP for treatment had significantly higher effect sizes ( $d = 0.43$ ;  $NNT = 4.20$ ) than studies in which patients were recruited through systematic screening ( $d = 0.13$ , not significantly different from zero;  $NNT = 13.51$ ).

### Conclusions

Although the number of studies was relatively low and the quality varied, psychological treatment of depression was found to be effective in primary care, especially when GPs refer patients with depression for treatment.

### Keywords

cognitive behaviour therapy; depression; major depression; meta-analysis; psychological treatment; psychotherapy.

## BACKGROUND

Depressive disorders are highly prevalent<sup>1,2</sup> and have a high incidence.<sup>3</sup> They are also associated with huge losses of quality of life in patients and their relatives,<sup>4,5</sup> increased mortality rates,<sup>6</sup> high levels of service use, and enormous economic costs.<sup>7-9</sup> Major depression is currently the fourth disorder worldwide in terms of disease burden, and is expected to be the disorder with the highest disease burden in high-income countries by the year 2030.<sup>10</sup>

Most depressive disorders are treated in primary care.<sup>11</sup> Although many GPs tend to prescribe antidepressant medications, the majority of patients prefer psychological treatments.<sup>12</sup> In the last few decades, several randomised studies have examined the effectiveness of psychological treatments in primary care patients.<sup>13,14</sup> Some of these found positive effects,<sup>15-17</sup> but several others found no significant effects.<sup>18-20</sup>

It is important to assess the overall effectiveness of treatments and to study possible determinants of

*P Cuijpers, PhD, professor of clinical psychology; A van Straten, PhD, associate professor of clinical psychology, Department of Clinical Psychology, VU University Amsterdam, and EMGO Institute, VU University Medical Center, Amsterdam; A van Schaik, MD, PhD, psychiatrist, EMGO Institute, VU University Medical Center, and Department of Psychiatry, VU University Medical Center, Amsterdam, the Netherlands. G Andersson, PhD, professor of clinical psychology, Department of Behavioural Sciences and Learning, Linköping University, Sweden and Department of Clinical Neuroscience, Psychiatry Section, Karolinska Institutet, Stockholm, Sweden.*

### Address for correspondence

Pim Cuijpers, Professor of Clinical Psychology, Department of Clinical Psychology, VU University Amsterdam, Van der Boerhorststraat 1, 1081 BT Amsterdam, the Netherlands. E-mail: p.cuijpers@psy.vu.nl

**Submitted:** 1 October 2008; **Editor's response:**

28 November 2008; **final acceptance:** 16 December 2008.

©British Journal of General Practice.

This is the full-length article of an abridged version published in print. Cite this article as: *Br J Gen Pract* 2009; DOI: 10.3399/bjgp09X395139.

treatment outcome, as psychological treatments for depression are widely applied in primary care. Patients enrolled in primary care generally have a less progressed developmental stage of their illness than patients in specialised healthcare settings,<sup>21,22</sup> and this may be related to the effectiveness of psychological treatments. However, only one earlier meta-analysis has integrated the results of studies on psychological treatments of depression in primary care.<sup>23</sup> Although this meta-analysis found evidence that psychological treatments are effective in primary care, it has several limitations. It included only a limited number of the currently available studies and did not conduct elaborate subgroup analyses to examine the heterogeneity of outcomes. Furthermore, by focusing only on studies in primary care, it does not answer the question whether these treatments are less effective than in other settings, as has been suggested in the literature.<sup>24</sup> Therefore, the aim of the present systematic review is to conduct a meta-analysis of randomised studies examining psychological treatments in primary care patients.

Several of these studies found small or even negative effects for these treatments, which led to suggestions that psychological treatments are less effective in primary care than in other settings.<sup>24</sup> As a consequence, multivariate meta-regression techniques were used on a large database of randomised controlled studies of psychological treatments of depression to examine whether these treatments are less effective when delivered in primary care as compared to other settings.

## METHOD

### *Identification and selection of studies*

A large database of studies of the psychological treatment of depression in general was used. This database, how it was developed, and the methods used have been described in detail elsewhere.<sup>25</sup> Key materials, overviews of the goals and mission, and an overview of all other published meta-analyses that have used this database can be downloaded from the website for this project ([www.psychotherapyrcrts.org](http://www.psychotherapyrcrts.org)). In brief, the database was developed through a comprehensive literature search (from 1966 to December 2007) in which 8861 abstracts in PubMed (1403 abstracts), PsycINFO (2097), EMBASE (2207), and the Cochrane Central Register of Controlled Trials (2204) were examined. Moreover, the Dissertation Abstracts International (950 abstracts) was examined to identify unpublished studies. These abstracts were identified by combining terms indicative of psychological treatment and depression (both MeSH terms and text words). In addition, the primary studies from 22 meta-analyses of psychological treatment of depression<sup>26</sup> were

## *How this fits in*

Psychological treatment of depression is effective in primary care patients. This meta-analysis found no indication that psychological treatments of depression are less effective in primary care than in other settings, as long as the patients are referred by their GP. Studies in which patients were recruited through systematic screening seem to result in lower effect sizes.

collected and the references of included studies were examined. A total of 857 published papers and 33 dissertations were retrieved for further study (Figure 1).

The abstracts from the bibliographic databases were screened and the retrieved reports were examined independently by two reviewers for possible inclusion. When the two reviewers disagreed, they discussed the differences with the third reviewer until agreement was reached.

This meta-analysis included: (a) randomised controlled trials (b) that examined adult primary care patients (c) with a depressive disorder or an elevated level of depressive symptomatology (d) in which the effects of a psychological treatment (e) delivered in primary care (f) was compared to a control condition. Primary care patients had been recruited through direct referrals from GPs, or through systematic screening of patients waiting to see their GP or who had recently been seen by the doctor.

The same inclusion criteria were used for the meta-regression analysis, in which psychological treatment in primary care was compared with other settings, except that all studies on any adult patient group were included.

Psychological treatments were defined as interventions in which verbal communication between a therapist and a client was the core element, or in which a psychological treatment was written down in book format or a computer program (guided self-help or bibliotherapy) that the client worked through more or less independently, but with some kind of personal support from a therapist (by telephone, email, or otherwise). The following studies were excluded: studies on children and adolescents (<18 years of age); studies in which the psychological intervention could not be discerned from other elements of the intervention (managed care interventions and disease management programmes); studies in which a standardised effect size could not be calculated (mostly because no test was performed in which the difference between the experimental and the control groups was examined); studies on inpatients; studies focusing on maintenance treatments and relapse prevention; and studies that included participants who were either anxious or depressed (studies with patients

who were both anxious and depressed were included). Comorbid general medical or psychiatric disorders were not used as an exclusion criterion. No language restrictions were applied.

#### Data extraction

Studies were coded on three domains. First to be coded were several patient characteristics: target population (adults, older adults, specific population, such as women with postnatal depression, people with general medical disorders, and minority groups); recruitment method (open community recruitment, recruitment from clinical samples, recruitment in primary care, recruitment through systematic screening not in primary care, and other recruitment method); definition of depression (major depressive disorder diagnosed with a formal diagnostic interview such as the Composite International Diagnostic Interview or the Structural Clinical Interview, other definition of depression, usually depression defined as scoring above a cut-off score on a self-report scale, or depressive disorders including dysthymia, minor depression, or adjustment disorders). The severity of depression at baseline was investigated by examining the score on the Beck Depression Inventory (BDI)<sup>27</sup> at pre-test, as well as the score on the Hamilton Rating Scale for Depression (HAM-D),<sup>28</sup> although these data were only available for a limited number of studies.

Second to be coded were characteristics of the intervention (cognitive behaviour therapy, problem-solving therapy, and other therapies); format of the therapy (individual, group, and guided self-help); number of treatment sessions ( $\leq 6$  or  $\geq 7$ ). The professional background of the therapists (Table 1) was also coded, but was not examined further because of the wide diversity of these professional backgrounds.

Third to be coded were general characteristics of the design of the studies: type of control group (waiting list, care-as-usual, and other control group); type of analyses (intention-to-treat analyses, and completers-only analyses); and country (US, UK, and other).

#### Quality assessment

The validity of the studies was assessed using three basic criteria:<sup>29</sup> allocation to conditions was done by an independent (third) party; blinding of assessors of outcomes; and completeness of follow-up data. A fourth basic criterion (blinding of participants about the condition they were assigned to) was not used, because this is usually not possible in studies examining psychological treatments.

#### Analyses

Effect sizes (standardised mean difference, *d*) were calculated for each study by subtracting (at post-test) the average score of the control group from the average score of the experimental group and dividing the result by the pooled standard deviations (SDs) of the experimental and control groups. Thus, an effect size of 0.5 indicates that the mean of the experimental group is half an SD larger than the mean of the control group. Effect sizes of 0.80 and higher can be assumed to be large, while effect sizes of 0.50 to 0.80 are moderate, and lower effect sizes are considered to be small.<sup>30</sup>

Only those instruments that explicitly measure symptoms of depression were used in the calculations of effect sizes (Table 1). If more than one measurement of depression had been used, the mean of the effect sizes was calculated, so that each study (or contrast group) only contributed one effect size. When means and SDs were not reported, other statistics (*t*-value, *P*-value) were used to calculate effect sizes.

The standardised mean difference is not easy to interpret from a clinical point of view and so the numbers-needed-to-treat (NNT) were also calculated, using the formulae provided by Kraemer.<sup>31</sup> The NNT is defined as the number of patients one would expect to treat with a psychological treatment to have one more successful outcome compared to the same number of patients in the control group.

The Comprehensive Meta-analysis computer program (version 2.2.021), which was developed for support in meta-analysis, was used to calculate pooled mean effect sizes, and the random effects model<sup>29</sup> was used to conduct all analyses because considerable heterogeneity was expected.

The *Q* statistic was calculated as an indicator of heterogeneity. The *I*<sup>2</sup> statistic, which is an indicator of heterogeneity in percentages, was also calculated. A value of 0% indicates no observed heterogeneity, and larger values show increasing heterogeneity, with 25% as low, 50% as moderate, and 75% as high heterogeneity.<sup>32</sup>

Publication bias was tested by inspecting the funnel plot of the meta-analysis, and by using Duval and Tweedie's trim and fill procedure,<sup>33</sup> which yields an estimate of the effect size after publication bias has been taken into account. This procedure is based on the expectation that if no publication bias is present the effect sizes will be dispersed equally on either side of the overall effect. An asymmetric funnel plot can be seen as an indication for publication bias (if there is a relatively large number of small studies falling towards the right of the mean effect and relatively few falling towards the left). Duval and Tweedie developed a method that allows imputation of missing studies. This method

**Table 1. Selected characteristics of randomised controlled studies examining psychological treatment in primary care.**

Study	Recruitment	Definition of depression	Psychological treatment	Sessions <i>n</i>	<i>n</i>	Format	Treatment provider	Control	<i>n</i>	Measurements	Instruments	ITT	Country
Barrett <i>et al</i> , 2001 <sup>18</sup>	NR	DYS or minD + HAM-D >10	PST	80	6	Ind	Psychol	Placebo	81	Pre, post	HSCL-D-20	+	US
Conradi <i>et al</i> , 2007 <sup>34</sup>	Referral	MDD (CIDI)	CBT (+PE)	44	14	Ind	Psychol	CAU	72	Pre, post, 3, 6, 36 mo	BDI	+	NL
King <i>et al</i> , 2000 <sup>35</sup>	Referral	BDI >14 + depressed according to GP	CBT Non-directive counseling	63 67	6	Ind	Psychol Psychol	CAU	67	Pre, post, 12 mo	BDI	+	UK
Lynch <i>et al</i> , 2004 <sup>19</sup>	Screening	HAM-D 11–26, sD (High MOS-DSI; no MDD/DYS)	PST Stress management	9 18	6, t 6, t	Ind	Nurse Nurse	CAU	13	Pre, Post	BDI, DHP-D	–	US
Lynch <i>et al</i> , 1997 <sup>36</sup>	Screening	sD (High MOS-DSI; no MDD/DYS)	PST	11	6	Ind	Students	CAU	13	Pre, Post	HAM-D, BDI	–	US
Mynors-Wallis <i>et al</i> , 1995 <sup>15</sup>	Referral	MDD (RDC) + HAM-D >13	PST	30	6	Ind	Psychiatr + GPs	Placebo	30	Pre, post	HAM-D, BDI	+	UK
Ross & Scott, 1985 <sup>16</sup>	Referral	MDD (PSE/RDC) + BDI >15	CBT (Ind) CBT (group)	21 9	12 12	Ind Grp	Social worker Social worker	WL	21	Pre, post, 3, 6, 12 mo	BDI, MADRS	–	UK
Schulberg <i>et al</i> , 1996 <sup>17</sup>	Screening	MDD (DIS/DSM-III-R) + HAM-D >13	IPT	93	16	Ind	Psychiatr + psychol	CAU	92	Pre, post, 2, 4 mo	HAM-D	+	US
Scott & Freeman, 1992 <sup>38</sup>	Referral	MDD (DSM-III)	CBT Counseling	29 29	16 16	Ind Ind	Psychol Social workers	CAU	29	Pre, post	HAM-D	–	UK
Scott <i>et al</i> , 1997 <sup>39</sup>	Referral	MDD (DSM-III-R) + BDI >20	CBT	18	6	Ind	CBT therapist	CAU	16	Pre, post, 3, 6, 12 mo	HAM-D, BDI	–	UK
Scott & Stradling, 1990 <sup>37</sup>	Referral	Depressive disorder (RDC/PSE) + BDI >14	CBT (ind) CBT (grp)	27 17	12 12	Ind Grp	NR NR	WL	23	Pre, post, 3, 6, 12 mo	BDI, MADRS, IDAQ	–	UK
Simpson <i>et al</i> , 2003 <sup>20</sup>	Screening	BDI 14–40, depressed for 6 months	Psychodynamic counseling	73	6–12	Ind	Counselors	CAU	72	Pre, post, 12 mo	BDI	+	UK
Teasdale <i>et al</i> , 1984 <sup>40</sup>	NR	MDD (RDC) + BDI >20 + HAM-D >14	CBT	17	15	Ind	Psychol	CAU	17	Pre, post, 3 mo	MADRS, BDI, HAM-D	+	UK
Van Schaik <i>et al</i> , 2006 <sup>24</sup>	Screening	GDS-15 >5 + MDD (Prime-MD)	IPT	69	10	Ind	Psycholog + psychiatr nurses	CAU	74	Pre, post	MADRS, GDS	+	NL
Williams <i>et al</i> , 2000 <sup>41</sup>	Screening	DYS or minD + HAM-D >10	PST	80	6	Ind	Psycholog + social workers + counselors	Placebo	81	Pre, Post	HSCL-D-20	+	US

BDI = Beck Depression Inventory. CAU = care-as-usual. CBT = cognitive-behavioural therapy. CIDI = Composite International Diagnostic Interview. DHP-D = Duke Health Profile – depression scale. DIS = Diagnostic Interview Schedule. DSM-III-R = The Diagnostic and Statistical Manual of Mental Disorders – third revision. DYS = dysthymia. GDS = Geriatric Depression Scale. Grp = group therapy. HAM-D = Hamilton Rating Scale for Depression. HSCL-D-20 = Hopkins Symptom Checklist Depression Scale. IDAQ = Irritability, Depression, Anxiety Questionnaire. Ind = individual therapy. IPT = interpersonal psychotherapy. ITT = intention-to-treat. MADRS = Montgomery-Åsberg Depression Rating Scale. MDD = major depressive disorder. minD = minor depression. Mo = months. MOS-DSI = Medical Outcomes Study – Depression Screening Inventory. NL = the Netherlands. NR = not reported. PE = psycho-education. Prime-MD = The Primary Care Evaluation of Mental Disorders. PSE = Present State Examination. PST = problem-solving therapy. psychiatr nurse = psychiatric nurse. Psychiatr = psychiatrist. Psychol = psychologist. RDC = Research Diagnostic Criteria. sD = subthreshold depression. t = telephone sessions. WL = waiting-list control group.

determines where the missing studies are likely to fall, adds them to the analysis, and then recomputes the combined effect.

Subgroup analyses and univariate meta-regression analyses were conducted according to the procedures implemented in Comprehensive Meta-analysis (version 2.2.021). The subgroup-analyses used mixed-effects analyses that pooled studies within subgroups with the random effects model, but tested for significant differences between subgroups with the fixed effects model.

Multivariate meta-regression analyses in which

more than one predictor was entered simultaneously were conducted in STATA/SE (8.2 for Windows), because these analyses cannot be conducted in Comprehensive Meta-analysis. To avoid collinearity among the predictors that were entered in the regression models, it was examined whether high correlations were found among the variables that could be entered into the model. The correlations between all variables described in the 'Data extraction' section above were calculated and it was checked whether the correlations were lower than  $r = 0.60$ .

## RESULTS

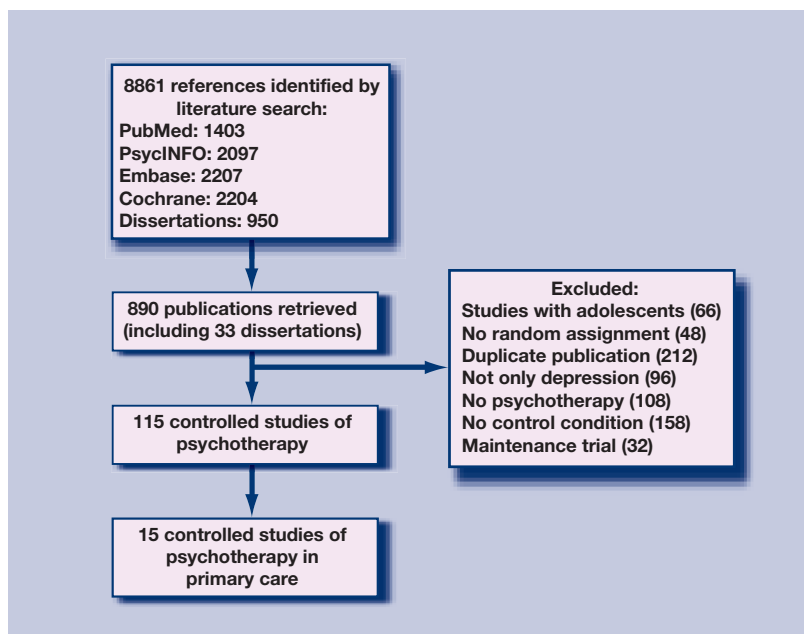
### Description of included studies

All inclusion criteria for studies on psychological treatment in primary care were met by 15 studies, in which 20 psychological treatment conditions were compared to a control group.<sup>15-21,34-41</sup> In these 15 studies, a total of 1505 patients participated (804 in the psychological treatment conditions and 701 in the control conditions). Selected characteristics of the studies included are presented in Table 1. A flow chart of the inclusion of studies and the reasons for exclusion are presented in Figure 1.

Patients in seven studies were referred to the study by the GP, while in six other studies patients were screened for depression (while waiting for their GP, or by a postal questionnaire sent to patients who had recently been seen by their GP). Mixed methods were used in the remaining two studies, or the recruitment method was not clear. Other characteristics of the studies and the comparisons between psychotherapies and control conditions can be found in Table 2 (diagnosis; type of psychotherapy; number of sessions; type of control group, intention-to-treat analyses, and country where the study was conducted).

The pre-test scores on the BDI were available in eight studies (11 comparisons). In 10 of these 11 comparisons the pre-test BDI scores fell inside the range of moderate to severe depression (BDI between 19 and 29), while the mean BDI score was in the severe range at pre-test in only one of the studies. The pre-test score on the HAM-D was available in seven studies (eight comparisons) and ranged from 13.3 to 22.3 (six comparisons below 20 and two above).

Figure 1. Flowchart of inclusion of studies.



The quality of the studies varied. Blinding of assessors was reported in 13 studies and eight studies reported that allocation to conditions had been conducted by an independent party. Drop-out numbers ranged from 3.3% to 41.2% (one study did not report drop-out). Intention-to-treat analyses were conducted in 10 studies and all patients who were randomised were used in these analyses, whether or not they dropped out of the intervention or study; the other studies were limited to completers-only analyses.

### Psychological treatment in primary care versus control at post-test

The overall mean effect size of psychological treatment versus the control conditions at post-test was 0.31 (95% CI = 0.17 to 0.45), which is usually considered to be a small effect.<sup>30</sup> Heterogeneity was significant ( $Q = 34.91$ ;  $P < 0.05$ ), but low to moderate ( $I^2 = 45.58$ ). The NNT which corresponds to an effect size of 0.31 was 5.75. Details of these results are shown in Table 2 and Figure 2.

The analyses included studies in which more than two psychological treatments were compared to a control group, which means that multiple comparisons from one study were included in the same analysis. These multiple comparisons are not independent of each other, which may have resulted in an artificial reduction of heterogeneity and a bias in the overall mean effect size. Additional analyses were conducted as a consequence, in which only one comparison per study was included (Table 2). Only the comparison with the largest effect size was included first, followed by another analysis including only the smallest effect size. Table 2 shows that the results did not differ very much from those in which all comparisons were included.

Comparable effect sizes were found ( $d = 0.43$ ; 95% CI = 0.22 to 0.64;  $Q = 26.30$ , not significant;  $I^2 = 50.57$ ;  $n = 14$ ) when the analyses were limited to the effect sizes found for the BDI. The result was the same when the analyses were limited to the effect sizes found for the HAM-D, which is a clinical interview ( $d = 0.49$ ; 95% CI = 0.30 to 0.68;  $n = 7$ ); however, heterogeneity was zero in these analyses ( $Q = 5.34$ ; not significant;  $I^2 = 0$ ).

Neither the funnel plots nor Duval and Tweedie's trim and fill procedure indicated a significant publication bias. The effect size decreased somewhat after adjustment for possible publication bias (adjusted effect size:  $d = 0.20$ ; 95% CI = 0.05 to 0.34), but the observed and adjusted effect size did not differ significantly.

The effects of psychological treatments at follow-up were not examined. No effect sizes were available at follow-up in nine studies (either because no



**Table 2. Meta-analyses of studies examining effects of psychological treatment of depression in primary care compared to control conditions at post-test: overall results and subgroup analyses.<sup>a</sup>**

	<i>n</i> <sub>studies</sub>	<i>n</i> <sub>comp</sub>	<i>d</i>	95% CI	Z	Q	<i>I</i> <sup>2</sup>	<i>P</i> -value <sup>g</sup>	NNT
All studies	20	20	0.31	0.17 to 0.45	4.25 <sup>d</sup>	34.91 <sup>b</sup>	45.58		5.75
One effect per study (highest)	15	15	0.30	0.13 to 0.48	3.47 <sup>c</sup>	30.35 <sup>c</sup>	53.88		5.95
One effect per study (lowest)	15	15	0.25	0.09 to 0.41	3.10 <sup>c</sup>	27.08 <sup>b</sup>	48.30		7.14
BDI only	10	14	0.43	0.22 to 0.64	4.07 <sup>d</sup>	26.30 <sup>b</sup>	50.57		4.20
HAM-D only	7	7	0.49	0.30 to 0.68	5.11 <sup>d</sup>	5.34, ns	0		3.68
Subgroup analyses									
Recruitment <sup>f</sup>									
Referral	7	11	0.43	0.28 to 0.58	5.75 <sup>d</sup>	9.71, ns	0	<sup>b</sup>	4.20
Screening	6	7	0.13	-0.08 to 0.34	1.23	9.82, ns	38.89		13.51
Diagnosis									
MDD	7	10	0.21	-0.00 to 0.42	1.95 <sup>e</sup>	19.15 <sup>b</sup>	53.00	ns	8.47
Other	7	10	0.40	0.23 to 0.56	4.61 <sup>d</sup>	11.12, ns	19.09		4.50
Type of treatment <sup>h</sup>									
CBT	7	9	0.42	0.22 to 0.62	4.16 <sup>d</sup>	9.72, ns	17.67	ns	4.27
PST	5	5	0.19	-0.15 to 0.53	1.08, ns	9.77 <sup>b</sup>	59.05		9.43
Other	5	6	0.27	0.04 to 0.49	2.35 <sup>b</sup>	9.65 <sup>e</sup>	48.20		6.58
Number of sessions <sup>h</sup>									
≤6	7	9	0.25	0.02 to 0.48	2.12 <sup>b</sup>	16.56 <sup>b</sup>	51.68	ns	7.14
≥7	8	11	0.36	0.17 to 0.54	3.75 <sup>d</sup>	17.34 <sup>e</sup>	42.34		5.00
Control group									
Care-as-usual	10	13	0.29	0.14 to 0.43	3.87 <sup>d</sup>	15.88, ns	24.45	ns	6.17
Other	5	7	0.40	0.07 to 0.73	2.39 <sup>b</sup>	18.73 <sup>c</sup>	67.97		4.50
Analyses									
Intention-to-treat	10	10	0.23	0.05 to 0.42	2.47 <sup>b</sup>	23.18 <sup>c</sup>	61.18	<sup>e</sup>	7.69
Completers-only	5	10	0.47	0.27 to 0.68	4.48 <sup>d</sup>	6.53, ns	0		3.85
Country									
UK	8	12	0.45	0.28 to 0.62	5.19 <sup>d</sup>	15.09, ns	27.09	<sup>b</sup>	4.00
US	5	6	0.11	-0.15 to 0.38	0.84, ns	10.01 <sup>e</sup>	50.03		16.13
NL	2	2	0.10	-0.15 to 0.35	0.81, ns	0.50, ns	0		17.86

<sup>a</sup>According to the random effects model. <sup>b</sup>*P*<0.05. <sup>c</sup>*P*<0.01. <sup>d</sup>*P*<0.001. <sup>e</sup>*P*<0.1. <sup>f</sup>In two studies information about how patients were screened was insufficient.

<sup>g</sup>The *P*-value indicates whether the difference between subgroups is significant. <sup>h</sup>The total number of studies is larger than 15 because some studies examined more than one type of psychotherapy. *n*<sub>comp</sub> = number of comparisons. *d* = standardised mean difference. *Q* = indicator of homogeneity. *I*<sup>2</sup> = indicator of heterogeneity in percentages. NNT = numbers-needed-to-treat. BDI = Beck Depression Inventory. HAM-D = Hamilton Rating Scale for Depression. MDD = major depressive disorder. CBT = cognitive behavioural therapy. PST = problem solving therapy. NL = the Netherlands.

follow-up assessment took place, or because a waiting list control group was used and this control group had received treatment at follow-up). The attrition rate was higher than 50% in one of the remaining six studies,<sup>39</sup> while the follow-up periods ranged from 3 to 36 months in the other five studies. Because of the small number of studies and the large differences in follow-up periods, the results of the treatments at follow-up were not pooled.

### Subgroup analyses

Several subgroup analyses were conducted, using the characteristics of the studies as described in the above section on 'Data extraction', except that subgroup analyses were not conducted with the target population, because all studies focused on adults in general. In addition, recruitment method (community recruitment, recruitment from clinical samples, and other recruitment methods) was not examined in these subgroup analyses, because patients in all these

studies were recruited through primary care. The results of the subgroup analyses are presented in Table 2.

The studies in which patients were recruited through systematic screening resulted in a significantly lower effect size (*d* = 0.13; 95% CI = -0.08 to 0.34; *Q* = 9.82, not significant; *I*<sup>2</sup> = 38.89) than studies in which patients were referred directly by the GP (*d* = 0.43; 95% CI = 0.28 to 0.58; *Q* = 9.71, not significant; *I*<sup>2</sup> = 0). The effect size for the studies in which patients were recruited through systematic screening was not significantly different from zero. There was a significant association (*P*<0.05) between effect size and the country where the study was conducted. Studies in the UK found a higher mean effect size (*d* = 0.45; 95% CI = 0.28 to 0.62) than studies in the US (*d* = 0.11; 95% CI = -0.15 to 0.38) and the two studies in the Netherlands (*d* = 0.10; 95% CI = -0.15 to 0.35); heterogeneity was low to moderate in the three subgroups, and the results of the studies in the US

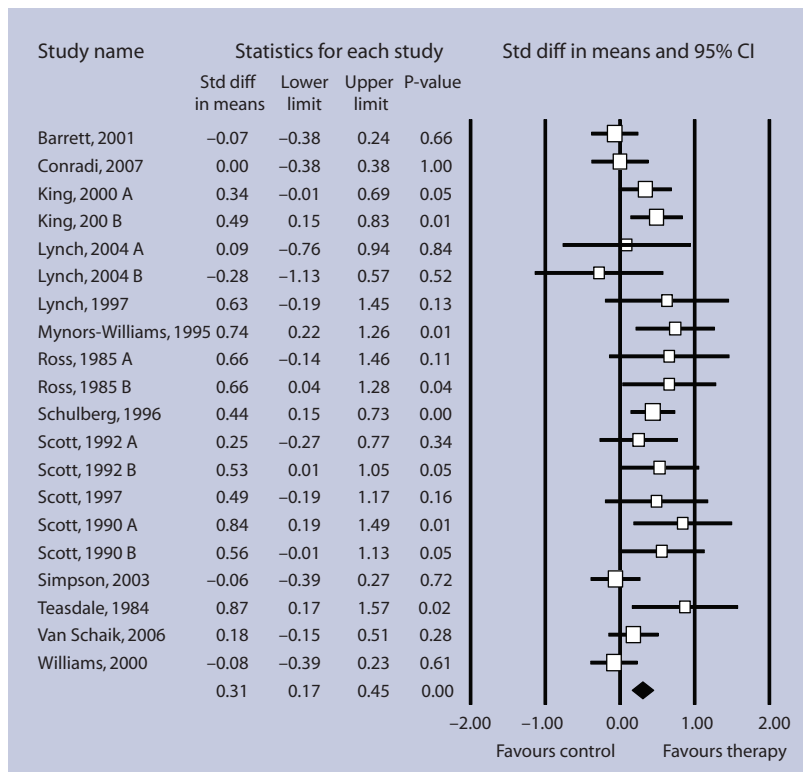


Figure 2. Standardised effect sizes of psychological treatment of depression in primary care patients compared to control conditions at post-test (A and B indicate two treatments from one study that were compared to a control group).

and the Netherlands were not significantly different from zero. There was a trend ( $P < 0.1$ ) indicating that studies using intention-to-treat analyses resulted in lower effect sizes ( $d = 0.23$ ; 95% CI = 0.05 to 0.42) than studies using completers-only analyses ( $d = 0.47$ ; 95% CI = 0.27 to 0.68).

Most BDI and HAM-D scores at pre-test were in the same range of severity (moderate to severe), which meant it was not possible to examine whether severity at pre-test was related to the effect size.

**Psychological treatment in primary care versus other settings**

The effect sizes found for psychological treatments in primary care were relatively small compared to the results of psychological treatments for depression in

general<sup>42,43</sup> and, therefore, treatments delivered in primary care were compared directly with these treatments when conducted in other settings (such as in samples recruited through media announcements, or in clinical samples from specialised mental health care). A literature search identified 99 studies in which psychological treatments in other outpatient settings than primary care were compared to a control condition. These 99 studies included a total of 154 comparisons between a psychological treatment and a control group; 6427 patients with depression participated in these studies: 3843 in the experimental groups and 2584 in the control groups. Patients in 55 of the 99 studies (55.6%) were recruited from the community; patients in nine studies (9.1%) were recruited from specialised mental health services; patients in 18 (18.2%) studies were recruited through systematic screening (not in primary care, but in community or general medical samples); and 17 studies (17.2%) used other recruitment methods. Other selected characteristics of the 154 comparisons can be found on the website of this project ([www.psychotherapyprcts.org](http://www.psychotherapyprcts.org)).

The effects of the psychological treatments in primary care ( $d = 0.31$ ; 95% CI = 0.17 to 0.45) were compared with those in other settings ( $d = 0.67$ ; 95% CI = 0.58 to 0.75), and it was found that this difference was highly significant ( $P < 0.001$ ). Results of these analyses are presented in Table 3. Very high effect sizes ( $d > 2.0$ ) were found in some studies (conducted in settings other than primary care) and the possibility that these could be outliers led to another analysis from which these potential outliers were excluded. The difference between studies in primary care and those in other settings remained highly significant ( $P < 0.001$ ) in these analyses. This difference was also highly significant when only one effect size per study was used and when the analyses were limited to effect sizes based on the HAM-D and the BDI (Table 3).

**Multivariate meta-regression analyses**

Table 3. A meta-analytic comparison of psychological treatment of depression in primary care and other settings

	Studies in primary care <sup>a</sup>				Studies in other settings							
	n	d	95% CI	NNT	n	d	95% CI	Z	Q	I <sup>2</sup>	NNT	P-value
All studies	20	0.31	0.17 to 0.45	5.75	154	0.75	0.65 to 0.84	15.49	519.52	70.55	2.48	<0.001
Outliers excluded	20	0.31	0.17 to 0.45	5.75	146	0.67	0.58 to 0.75	15.34	388.29	62.66	2.75	<0.001
One effect size per study (highest)	15	0.30	0.13 to 0.48	5.95	99	0.79	0.67 to 0.91	12.59	386.57	74.65	2.36	<0.001
One effect size per study (lowest)	15	0.25	0.09 to 0.41	7.14	99	0.62	0.51 to 0.73	11.38	297.08	67.01	2.96	<0.001
BDI only	14	0.43	0.22 to 0.64	4.20	98	0.80	0.68 to 0.92	12.96	288.16	66.34	2.34	<0.01
HAM-D only	7	0.49	0.30 to 0.68	3.68	50	0.97	0.79 to 1.15	10.77	149.85	67.30	1.97	<0.001

<sup>a</sup>More details about these effect sizes can be found in Table 2. BDI = Beck Depression Inventory. d = standardised mean difference. HAM-D = Hamilton Rating Scale for Depression. I<sup>2</sup> = indicator of heterogeneity in percentages. NNT = numbers-needed-to-treat. Q = indicator of homogeneity.

As the significant difference between studies in primary care and those in other settings could have been influenced by characteristics of the populations, the interventions, and the design of the studies, multivariate meta-regression analyses were conducted which controlled for these variables.

For the purposes of the first meta-regression analysis (with the effect size as the dependent variable), all characteristics described in the above 'Data extraction' section were entered as predictors, after transforming them into dummy variables. Recruitment through primary care was one of the predictors in these analyses, the results of which are presented in Table 4, and it is clear that recruitment in primary care was a significant predictor of the effect size, even after controlling for all other characteristics of the studies ( $B = -0.41$ ; standard error [SE] = 0.19;  $P = 0.03$ ).

Indications that there were two types of recruitment in primary care resulted in the same meta-regression analysis being conducted once more (with the same predictors), but this time two separate dummy variables were entered for recruitment in primary care: the first one indicated referral by the GP to the treatment, and the second one indicated systematic screening of primary care patients. Results of these analyses are presented in Table 4 and show that referral by the GP was not a significant predictor ( $P = 0.38$ ), but systematic screening was ( $B = -0.60$ ; SE = 0.23;  $P = 0.01$ ).

## DISCUSSION

### Summary of main findings

The psychological treatment of depression is effective in primary care patients. Although the effects seemed to be lower than the effect sizes found for psychological treatments in other settings, these findings suggest that the studies in which patients from primary care populations were screened systematically were responsible for this lower effect size. When patients are referred by their GP for psychological treatment, no indication that the effects are lower than the effect sizes found for psychological treatment of depression in other settings was found. Earlier systematic reviews found indications that screening for depression in healthcare settings results in positive outcomes.<sup>44</sup> However, in the current study no evidence was found that psychological treatment is effective if patients are recruited through systematic screening in primary care, although this finding should be considered with caution because of the limitations of this study.

### Strengths and limitations of the study

This study has several limitations. First of all, the

**Table 4. Regression coefficients of study characteristics in relation to the effect size of psychological interventions for depression: multivariate meta-regression analyses with recruitment in primary care as predictor.**

Variable	$n_{comp}$	Full model			Full model (with screening and referral as separate predictors) <sup>a</sup>			
		B	SE	$P$ -value	B	SE	$P$ -value	
<b>Recruitment method</b>								
Community	96	Ref			Ref			
Primary Care	20	-0.41	0.19	0.03	Screening <sup>a</sup>	-0.60	0.23	0.01
					Referral <sup>b</sup>	-0.21	0.24	0.38
Screening <sup>b</sup>	26	-0.11	0.17	0.53		-0.03	0.19	0.88
Clinical	13	-0.02	0.19	0.93		-0.10	0.17	0.57
Other	19	-0.03	0.16	0.86		-0.05	0.16	0.38
<b>Target group</b>								
Adults	95	Ref			Ref			
Older adults	21	-0.10	0.14	0.47		-0.07	0.14	0.61
Specific population	58	0.08	0.12	0.50		0.09	0.12	0.47
Diagnosed MDD (y/n)	43	0.14	0.11	0.20		0.12	0.11	0.29
<b>Type of treatment</b>								
Other treatment	75	Ref			Ref			
CBT	85	0.02	0.10	0.84		0.00	0.10	0.99
PST	14	0.28	0.18	0.11		0.34	0.18	0.06
<b>Treatment format</b>								
Individual	87	Ref			Ref			
Group	68	-0.18	0.10	0.09		-0.19	0.10	0.06
Guided self-help	18	-0.14	0.18	0.45		-0.13	0.18	0.46
Number of sessions (<6 vs ≥7)	56	-0.14	0.11	0.22		-0.11	0.11	0.32
<b>Control group</b>								
Waiting list	97	Ref			Ref			
Care-as-usual	53	-0.38	0.13	0.004		-0.34	0.13	0.01
Other	24	-0.62	0.14	<0.001		-0.61	0.14	<0.001
Intention to treat analyses (y/n)	52	-0.06	0.11	0.58		-0.04	0.11	0.70
<b>Country</b>								
US	104	Ref			Ref			
UK	22	0.02	0.17	0.91		-0.09	0.18	0.60
Other	48	0.29	0.11	0.01		0.27	0.11	0.02
Constant	174	1.04	0.15	<0.001		1.03	0.15	<0.001

<sup>a</sup>In this model the studies in primary care were divided into those in which patients were recruited through systematic screening and those in which patients were referred.

<sup>b</sup>Systematic screening in populations, not in primary care. CBT = cognitive-behavioural therapy. MDD = major depressive disorder.  $n_{comp}$  = number of comparisons. PST = problem-solving therapy.

number of studies in primary care was relatively low, and the quality of the studies included was not optimal. Secondly, there were several important differences between these studies and the influence of these differences on the outcomes is not clear. Thirdly, there was considerable heterogeneity in most analyses, which suggests that the effects of therapies may be associated with, and perhaps confounded by, characteristics other than those examined in the subgroup analyses. A fourth limitation is that meta-regression techniques are known to have several limitations,<sup>45,46</sup> and their results should be considered with caution. Another important limitation is that many of the included studies compared psychological treatments with usual care. The



content of the 'usual care' arm in trials is crucial to interpreting the outcome of studies, but is often poorly described, not only in the papers reviewed, but also in other research fields.<sup>47</sup> However, usual care varies widely<sup>48</sup> and this may have distorted the results of the meta-analysis.

#### Comparison with existing literature

It is not clear why the effect sizes are so small when patients are recruited through systematic screening. Patients who do not actively seek treatment may have good reasons why they have not actively sought treatment before. One possibility is that their depression is less severe than patients in specialised mental health care and that the motivation to accept treatment is less strong. It is also possible that they think treatment will not be effective in their situation or expect their problems will disappear spontaneously. Another possibility is that those identified by screening are less severely or persistently depressed, and the screening instruments may not be subtle enough to detect such differences. Moreover, the possibility cannot be excluded that GPs inform and motivate patients to accept the offer of psychological treatment, whereas this happens to a much lesser extent in settings in which treatment recommendations follow a systematic screening procedure. Also, GPs implicitly use several other indicators to assess whether a patient will benefit from psychotherapy, such as former experience with psychotherapy of the patient, the physician's own experience with the patient in commitment with proposed treatments, and the role of specific temporary life events or somatic diseases that may resolve spontaneously. More research is needed to explore this finding.

#### Implications for future research and clinical practice

This study has several implications. First, GPs who refer to treatments and the treatment providers might need to collaborate closely as the right referral route appears to make a difference. Second, while screening tools may identify cases of depression, they might not screen effectively for suitability for psychological treatment. This could require that screening tools for suitability for psychological treatments should be developed.

Despite the limitations of this study, the results clearly suggest that psychological treatment is effective in primary care, but only when the patients with depression are referred by their GP for treatment. Systematic screening for patients with depression in primary care does not appear to lead to beneficial outcomes following psychological treatments.

#### Funding body

VU University Amsterdam, Linköping University and Karolinska Institutet

#### Competing interests

The authors have stated that there are none

#### Discuss this article

Contribute and read comments about this article on the Discussion Forum: <http://www.rcgp.org.uk/bjgp-discuss>

#### REFERENCES

- Kessler RC, McGonagle KA, Zhao S, *et al*. Lifetime and 12-months prevalence of DSM-III-R psychiatric disorders in the United States: results from the National Comorbidity Survey. *Arch Gen Psychiatry* 1994; **51**(1): 8–19.
- Alonso J, Angermeyer MC, Bernert S, *et al*. ESEMeD/MHEDEA 2000 Investigators, European Study of the Epidemiology of Mental Disorders (ESEMeD) Project. Prevalence of mental disorders in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. *Acta Psychiatr Scand Suppl* 2004; **(420)**: 21–27.
- Waraich P, Goldner EM, Somers JM, Hsu L. Prevalence and incidence studies of mood disorders: a systematic review of the literature. *Can J Psychiatry* 2004; **49**(2): 124–138.
- Ustun TB, Ayuso-Mateos JL, Chatterji S, *et al*. Global burden of depressive disorders in the year 2000. *Br J Psychiatry* 2004; **184**: 386–392.
- Saarni SI, Suvisaari J, Sintonen H, *et al*. Impact of psychiatric disorders on health-related quality of life: general population survey. *Br J Psychiatry* 2007; **190**: 326–332.
- Cuijpers P, Smit F. Excess mortality in depression: a meta-analysis of community studies. *J Affect Dis* 2002; **72**(3): 227–236.
- Berto P, D'Ilario D, Ruffo P, *et al*. Depression: cost-of-illness studies in the international literature, a review. *J Ment Health Policy Econ* 2000; **3**(1): 3–10.
- Greenberg PE, Birnbaum HG. The economic burden of depression in the US: societal and patient perspectives. *Exp Opin Pharmacother* 2005; **6**(3): 369–376.
- Smit F, Cuijpers P, Oostenbrink J, *et al*. Excess costs of common mental disorders: population-based cohort study. *J Ment Health Policy Econ* 2006; **9**(4): 193–200.
- Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. *PLoS Med* 2006; **3**(11): e442.
- Bijl RV, Ravelli A. Psychiatric morbidity, service use, and need for care in the general population: results of The Netherlands Mental Health Survey and Incidence Study. *Am J Public Health* 2000; **90**(4): 602–607.
- Schaik D, Klijn A, van Hout H, *et al*. Patients' preferences in the treatment of depressive disorder in primary care. *Gen Hosp Psychiatry* 2004; **26**(3): 184–189.
- Bower P, Rowland N. Effectiveness and cost effectiveness of counselling in primary care. *Cochrane Database Syst Rev* 2006; **3**: CD001025.
- Huibers MJ, Beurskens AJ, Bleijenberg G, van Schayck CP. Psychosocial interventions by general practitioners. *Cochrane Database Syst Rev* 2007; **3**: CD003494.
- Mynors-Wallis LM, Gath DH, Lloyd-Thomas AR, Tomlinson D. Randomised controlled trial comparing problem solving treatment with amitriptyline and placebo for major depression in primary care. *BMJ* 1995; **310**(6977): 441–445.
- Ross M, Scott M. An evaluation of the effectiveness of individual and group cognitive therapy in the treatment of depressed patients in an inner city health centre. *J R Coll Gen Pract* 1985; **35**(274): 239–242.
- Schulberg HC, Block MR, Madonia MJ, *et al*. Treating major depression in primary care practice. Eight-month clinical outcomes. *Arch Gen Psychiatry* 1996; **53**(10): 913–919.
- Barrett JE, Williams JW Jr, Oxman T E, *et al*. Treatment of dysthymia and minor depression in primary care: a randomized trial in patients aged 18 to 59 years. *Br J Fam Pract* 2001; **50**(5): 405–412.
- Lynch D, Tamburrino M, Nagel R, Smith MK. Telephone-based treatment for family practice patients with mild depression. *Psychol Rep* 2004; **94**(3 pt 1): 785–792.
- Simpson S, Corney R, Fitzgerald P, Beecham J. A randomized controlled trial to evaluate the effectiveness and cost-effectiveness of psychodynamic counselling for general practice patients with chronic depression. *Psychol Med* 2003; **33**: 229–239.
- Suh T, Gallo JJ. Symptom profiles of depression among general medical service users compared with specialty mental health service users. *Psychol Med* 1997; **27**(5): 1051–1063.

22. Cooper-Patrick L, Crum RM, Ford DE. Characteristics of patients with major depression who received care in general medical and specialty mental health settings. *Med Care* 1994; **32**(1): 15–24.
23. Bortolotti B, Menchetti M, Bellini F, *et al.* Psychological interventions for major depression in primary care: a meta-analytic review of randomized controlled trials. *Gen Hosp Psychiatry* 2008; **30**(4): 93–302.
24. Van Schaik A, van Marwijk H, Adèr H, *et al.* Interpersonal psychotherapy for elderly patients in primary care. *Am J Geriatr Psychiatry* 2006; **14**(9): 777–786.
25. Cuijpers P, van Straten A, Warmerdam L, Andersson G. Psychological treatment of depression: a meta-analytic database of randomized studies. *BMC Psychiatry* 2008; **8**: 36.
26. Cuijpers P, Dekker J. Psychologische behandeling van depressie: een systematisch overzicht van meta-analyses. *Ned Tijdschr Geneesk* 2005; **149**: 1892–1897.
27. Beck AT, Ward CH, Mendelson M, *et al.* An inventory for measuring depression. *Arch Gen Psychiatry* 1961; **4**: 561–571.
28. Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatry* 1960; **23**: 56–62.
29. Higgins JPT, Green S. Cochrane handbook for systematic reviews of interventions 4.2.5. In: *The Cochrane Library, 2005; Issue 3*. Chichester: John Wiley & Sons, Ltd. 2005.
30. Cohen J. *Statistical power analysis for the behavioral sciences*. Orlando, FL: Academic Press, 1977.
31. Kraemer HC, Kupfer DJ. Size of treatment effects and their importance to clinical research and practice. *Biol Psychiatry* 2006; **59**(11): 990–996.
32. Higgins JP, Thompson SG, Deeks JJ. Measuring inconsistency in meta-analyses. *BMJ* 2003; **327**(7414): 557–560.
33. Duval S, Tweedie R. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biomet* 2000; **56**(2): 455–463.
34. Conradi HJ, de Jonge P, Kluiters H, *et al.* Enhanced treatment for depression in primary care: long-term outcomes of a psycho-educational prevention program alone and enriched with psychiatric consultation or cognitive behavioral therapy. *Psychol Med* 2007; **37**(6): 849–862.
35. King M, Sibbald B, Ward E, *et al.* Randomised controlled trial of non-directive counselling, cognitive-behaviour therapy and usual general practitioner care in the management of depression as well as mixed anxiety and depression in primary care. *Health Technol Assess* 2000; **4**(19): 1–83.
36. Lynch DJ, Tamburrino MB, Nagel R. Telephone counseling for patients with minor depression: preliminary findings in a family practice setting. *J Fam Pract* 1997; **44**(3): 293–298.
37. Scott MJ, Stradling SG. Group cognitive therapy for depression produces clinically significant reliable change in community-based settings. *Behav Psychother* 1990; **18**(1): 1–19.
38. Scott AI, Freeman CP. Edinburgh primary care depression study: treatment outcome, patient satisfaction, and cost after 16 weeks. *BMJ* 1992; **304**(6831): 883–887.
39. Scott C, Tacchi MJ, Jones R, Scott J. Acute and one-year outcome of a randomised controlled trial of brief cognitive therapy for major depressive disorder in primary care. *Br J Psychiatry* 1997; **171**: 131–134.
40. Teasdale JD, Fennell MJ, Hibbert GA, Amies PL. Cognitive therapy for major depressive disorder in primary care. *Br J Psychiatry* 1984; **144**: 400–406.
41. Williams JW Jr, Oxman T, Frank E, *et al.* Treatment of dysthymia and minor depression in primary care: a randomized controlled trial in older adolescents. *JAMA* 2000; **284**(12): 1519–1526.
42. Churchill R, Hunot V, Corney R, *et al.* A systematic review of controlled trials of the effectiveness and cost-effectiveness of brief psychological treatments for depression. *Health Technol Assess* 2001; **5**(35): 1–173.
43. Cuijpers P, van Straten A, Warmerdam L, Smits N. Characteristics of effective psychological treatments of depression; a meta-regression analysis. *Psychother Res* 2008; **18**(2): 225–236.
44. Pignone MP, Gaynes BN, Rushton JL, *et al.* Screening for depression in adults: a summary of the evidence for the U.S. Preventive Services Task Force. *Ann Intern Med* 2002; **136**(10): 765–776.
45. Higgins JP, Thompson SG. Controlling the risk of spurious findings from meta-regression. *Stat Med* 2004; **23**(11): 1663–1682.
46. Thompson S, Higgins J. How should meta-regression analyses be undertaken and interpreted? *Stat Med* 2002; **21**(11): 1559–1573.
47. Somerville S, Hay E, Lewis M, *et al.* Content and outcome of usual primary care for back pain: a systematic review. *Br J Gen Pract* 2008; **58**(556): 790–797.
48. Mant D. The problem with usual care. *Br J Gen Pract* 2008; **58**(556): 755–756.