Psychological Treatment of Postpartum Depression: A Meta-Analysis

Pim Cuijpers, Jessica G. Brännmark, and Annemieke van Straten
VU University Amsterdam

Postpartum depression is a widespread and disruptive depressive disorder seriously affecting the lives of new mothers and their families. We conducted a meta-analysis of controlled and comparative studies of psychological treatments of postpartum depression. Seventeen studies were included. The mean standardized effect size of all psychological treatments compared to control conditions was 0.61 (95% CI: 0.37∼0.85). Several subgroup analyses were conducted. Studies with waiting list control groups had a larger mean effect size (0.96; 95% CI: 0.63∼1.29) than studies with a care-as-usual control group (0.41; 95% CI: 0.25∼0.58). No definite conclusions can be drawn about the longer term effects. Too few studies were available to draw conclusions about the relative effects of psychological treatments compared to pharmacological and other treatments. © 2007 Wiley Periodicals, Inc. J Clin Psychol 64: 103–118, 2008.

Keywords: postpartum depression; postnatal depression; psychological treatment; psychotherapy; meta-analysis.

Research in the last 20 years has shown Postpartum Depression (PPD) to be a widespread and disruptive depressive disorder affecting the lives of new mothers, their partners and children, and society at large (C. T. Beck, 1995; Lumley, Austin, & Mitchell, 2004; Murray & Cooper, 1997; Richards, 1990). PPD can be defined as a depressive disorder which occurs within 4 weeks' postpartum, although a time frame of 3 months after birth also has been suggested for defining PPD (Elliott, 2000). About one in every seven new mothers is affected by PPD (Wisner, Chambers, & Sit, 2006), resulting in an overall prevalence rate of 13% (M. W. O'Hara & Swain, 1996). Postpartum mood disorders represent the most frequent form of maternal morbidity following delivery (Stocky & Lynch, 2000).

Correspondence concerning this article should be addressed to: Dr. Pim Cuijpers, Professor of Clinical Psychology, Department of Clinical Psychology, VU University Amsterdam, Van der Boechorststraat 1, 1081 BT Amsterdam, The Netherlands; e-mail: P.Cuijpers@psy.vu.nl

JOURNAL OF CLINICAL PSYCHOLOGY, Vol. 64(1), 103–118 (2008)

Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/jclp.20432

Apart from the direct suffering caused by PPD in the patient and the increased risk of hospitalization (Dennis, 2004), several areas in the life of a patient can be adversely affected. PPD has been reported to result in an increased risk of marital stress and divorce (Holden, 1991), an increased risk of child abuse and neglect (Buist, 1998), and sometimes even in maternal suicide and infanticide (Sit, Rothschild, & Wisner, 2006). PPD also can have serious consequences for the children of affected mothers, in both the short and the long term (Murray & Cooper, 2004). The negative effects of maternal depression on children include an increased risk of impaired mental and motor development, difficult temperament, poor self-regulation, low self-esteem, and long-term behavioral problems (C. T. Beck, 1999; Goodman & Gotlib, 1999; Orvaschel, Walsh-Allis, & Ye, 1988; Wisner et al., 2006). It also can result in attachment insecurity (Hipwell, Goossens, Melhuish, & Kumar, 2000; Murray, 1992), social-interaction difficulties (Cummings & Davies, 1994; Dennis, 2004), and a negative influence on cognitive skills (Whiffen & Gotlib, 1989) and on expressive language development (Cox, Puckering, Pound, & Mills, 1987).

PPD often goes undetected due to lack of proper screening, and to the shame and loneliness that often make a woman hide it from her surroundings (Murray & Cooper, 1997). Untreated PPD often remits spontaneously after 4 to 6 months (M. H. O'Hara, 1997), but can in some cases easily last (much) longer, causing prolonged serious suffering (Cooper & Murray, 1998). Because it causes considerable distress and disruption to the women and their families, the delivery of effective treatment is generally considered a priority (Cooper, Murray, Wilson, & Romaniuk, 2003). To diagnose and begin treatment early, one could most likely prevent future suffering and disruption of the life of the individual, family relations, and the process of bonding and attachment with the baby.

In the past few decades, a number of studies have examined the efficacy of pharmacological and psychological treatments of PPD. Although some pharmacological interventions have been well-studied for depression unrelated to childbirth, methodological limitations render their efficacy equivocal for postpartum depression, with limited evidence available to guide practice or policy recommendations (Dennis, 2004). Furthermore, psychological interventions are usually preferred by mothers to antidepressant treatment, due to worries about safety issues with regard to breast-feeding (Appelby, Warner, Whitton, & Faragher, 1997).

Although several controlled trials have examined the effects of psychological treatments of PPD, the effects of these studies did not all point in the same direction. Although most studies found that psychological treatments were effective, some found strong and highly significant effects (Chabrol et al., 2002; Meager & Milgrom, 1996) while others found weak or even no effects (Fleming, Klein, & Corter, 1992; Prendergast & Austin, 2001). When the results of primary studies are not clear, a meta-analysis can be conducted to examine the true effect of these interventions. A meta-analysis also can be useful to examine why some studies find strong effects and others find only weak effects.

Several earlier meta-analyses and systematic reviews have examined the effects of psychological treatments of PPD. Some of these reviews focused mainly on the possibilities to predict and prevent the onset of PPD before the birth of the child (Dennis & Creedy, 2004; Richards, 1990). Two recent systematic reviews examined the effectiveness of psychological and other nonbiological treatments of PPD (Dennis, 2004; Gjerdingen, 2003). These reviews found some support for the effect of psychotherapy and home visits by a nurse; however, these reviews did not use meta-analytic methods to statistically integrate the results of the individual studies and

were not able to test whether different types of therapies differed significantly from each other.

In a recent study, the effects of nonpharmaceutical and nonhormonal interventions to reduce PPD were examined with a statistical meta-analysis (Lumley et al., 2004). In this study, clear indications were found that psychological interventions result in a consistent and substantial reduction of depressive symptoms. This finding was independent of the type of therapy and therapist background; however, this meta-analysis did not calculate standardized mean effect sizes (Cohen's *d*) for each study, but only relative risks of significant improvement. Improvement was typically defined as scoring below a cutoff score on a self-rating depression scale. Because of this dichotomization, much information was lost. Mean effect sizes are calculated in most meta-analyses of psychological treatments, thus the effects of treatments for PPD could not be compared to treatments of other depressive disorders. Furthermore, in the meta-analysis by Lumley and colleagues (2004), no subgroup analyses were conducted.

We therefore decided to conduct a new meta-analysis of psychological treatments of PPD, based on standardized mean effect sizes. In this study, we wanted to examine the effects of psychological treatments on PPD compared to control conditions and to other (nonpsychological) interventions.

Method

Search Strategy and Selection of Studies

Studies were traced through several methods. First, we used a large database of studies on the psychological treatment of depression in general, which has been described in detail elsewhere (Cuijpers, van Straten, Warmerdam, & Smits [in press], 2007a, b). This database was developed through a comprehensive literature search (from 1966 to March 2006) in which we examined a total of 4,661 abstracts in PubMed ($n = 1{,}127$ abstracts), PsychINFO ($n = 1{,}225$), EMBASE ($n = 925$), and the Cochrane Central Register of Controlled Trials ($n = 1{,}384$). We identified these abstracts by combining terms indicative of psychological treatment (i.e., psychotherapy, psychological treatment, cognitive therapy, behavior therapy, interpersonal therapy, reminiscence, life review) and depression (both MeSH-terms and text-words). For this database, we also collected the primary studies from 22 meta-analyses of psychological treatment of depression (Cuijpers & Dekker, 2005). We retrieved a total of 766 articles for further study. We examined these articles and selected the ones which focused on psychological treatments for PPD.

Second, we conducted additional searches in computerized literature databases, combining search terms indicative of postpartum depression (e.g., postpartum depression OR postpartum psychosis) and controlled trial (randomized OR randomized OR clinical OR trial OR experimental). Both keywords and text words were used. For these additional searches, we examined a total of 1,484 abstracts from EMBASE ($n = 395$ abstracts), the Cochrane database ($n = 58$), PubMed ($n = 142$), PsychINFO ($n = 751$), and Digital Dissertations ($n = 138$).

We also examined the references of earlier reviews on psychological treatment of PPD (Dennis, 2004; Dennis & Creedy, 2004; Gjerdingen, 2003; Lumley et al., 2004; Murray & Cooper, 1997; Richards, 1990), and we reviewed the reference lists of retrieved articles. We contacted the authors of studies that met inclusion criteria and asked whether they knew of any other (published and unpublished) studies in the field.

Studies were included if they compared the effects of a psychological intervention given to adult female participants with postpartum depression as diagnosed through clinical interview and/or self-report questionnaire to the effects of a control condition or another active intervention. No language restrictions were applied.

Our searches resulted in a total of 35 papers and reports (including two doctoral theses) on studies that possibly met our inclusion criteria. These papers were retrieved and studied. Seventeen studies (described in 21 papers) met our inclusion criteria. The remaining 14 studies were excluded because they did not use a control or comparison group ($n = 4$ studies), they examined preventive interventions and not treatment ($n = 3$), they examined nonpsychological interventions ($n = 3$), they were not aimed at postpartum depression ($n = 3$), and the effect size could not be calculated (and because no test examining the difference between experimental and control condition was conducted) ($n = 1$).

Four of the included studies examined social support interventions in which peers gave social support to women with PPD in group format ($n = 3$ studies) or in an individual format. Although it could be disputed whether this should be considered a psychological intervention, we decided to include these studies and examine in the meta-analyses whether the effect sizes found therein differed from the other studies. Selected characteristics of the included studies are presented in Table 1.

Results

Description of Studies

In the 17 included studies, there were 1,248 participants: 700 in the psychological treatment conditions, 460 in the control conditions, and 88 in other treatments. In 12 studies, participants were selected through systematic screening of women visiting health services; in the remaining 5 studies, the women were recruited through referrals and community recruitment. Participants had to meet diagnostic criteria of a (major or minor) depressive disorder in 8 studies whereas in the other studies only a high score on a self-report depression questionnaire was sufficient for participation. The Edinburgh Postnatal Depression Scale (EPDS; Cox, Holden, & Sagovsky, 1987) was used as a measure of depression in 14 studies while the Beck Depression Inventory (BDI; A. T. Beck, Ward, Mendelson, Mock, & Erbauch, 1961) was used in 7 studies. Cognitive-behavioral therapy was examined in 6 studies, social support interventions were examined in 5 studies, and 2 studies examined interpersonal psychotherapy. The intervention was delivered in an individual format in 10 studies while a group format was used in 5 studies (One used a combined group and individual format.) In 14 studies, a psychological intervention was compared to a control condition (care-as-usual in 9 studies, a waiting list in 3 studies, and another control group in 2 studies). Four studies were conducted in the United Kingdom, 4 in Australia, 3 in Canada, 2 in the United States, and the remaining 4 in other countries (i.e., France, Taiwan, Turkey, Sweden).

Quality Assessment

At least 25 scales are available to assess the validity and quality of randomized controlled trials (Higgins & Green, 2005); however, there is no evidence that these scales provide more reliable assessments of validity. We preferred, therefore, to use a

Table 1
Selected Characteristics of Randomized Controlled Trials Examining the Effects of Psychological Interventions for Women With Postpartum Depression

Study	Recruitment	Definition of depression[a]	Condition	N	Format	Intervention	Measure-ments	Instruments[a]	Country	Attr (%)	BA[b]	All	RA
Appleby, Warner, Whitton, & Faragher, 1997	Community-based SCR	EPDS > 10+ RCIS > 12 +MDD or minD (RDC)	1. CBT (6 s)+ PLA 2. CBT (6s)+ fluoxetine 3. CBT (1s)+ PLA 4. CBT (1s) +fluoxetine	21 21 23 22	Ind	Counseling with CBT, practical advice (6 s)	Pre, post	RCIS, EPDS	UK	30	+	+	+
Armstrong & Edwards, 2004	Referrals and community recruitment	EPDS ≥ 12	1. Support group 2. Pram-walking	10 9	Grp	Social support group (12 s)	Pre, post	EPDS	AU	21	–	+	+
Chabrol et al., 2002	SCR in obstetric clinic	EPDS ≥ 11	1. Counseling 2. CAU	21 38	Ind	support, education+CBT (5–8 s)	Pre, post	HDRS, BDI, EPDS	FR	20	–	–	–
Chen, Tseng, Chou, & Wang, 2000	SCR postnatal wards of hospital	BDI ≥ 9/10	1. Support group 2. CAU	30 30	Grp	Psychoeducation+mutual support (4–5 s)	Pre, post	BDI	Tai-wan	NR[c]	–	–	+
Clark, Wenzel, & Tluczek, 2003	Referrals and community recruitment	MDD (DSM-IV)	1. Mother–infant therapy group 2. IPT 3. WL	13 15 11	Grp	1. Support group+support children (12 s) 2. Interpersonal psychotherapy (12 s)	Pre, post	BDI, CES-D	US	18	+	–	–
Cooper, Murray, Wilson, & Romaniuk, 2003	SCR of primiparous women in hospital records	EPDS ≥ 12+ MDD (DSM-III-R/SCID)	1. CBT 2. Psychodynamic therapy	43 50 48 52	Ind	3 standardized treatments (10 s)	Pre, post, 9, 18, 60 mn	EPDS	UK	12	+	–	+

TABLE 1. Continued

Study	Recruitment	Definition of depression[a]	Condition	N	Format	Intervention	Measure-ments	Instruments[a]	Country	Attr (%)	BA[b]	All	RA
			3. Counseling 4. CAU										
Dennis, 2003	SCR in immunization clinics	EPDS ≥10	1. Peer support 2. CAU	20 22	Ind	Individual support by peers	Pre, post	EPDS	CAN	2	+	+	+
Fleming, Klein, & Corter, 1992	SCR at maternity ward of hospital	High score on CES-D +1 of 3 other scales	1. Support group 2. CAU 3. Group-by-mail	44 88 15	Grp	Unstructured mutual support group (8 s)	Pre, post	CES-D, EPDS, CES, MAACL	CAN	7	−	−	−
Holden, Sagovsky, & Cox, 1989	SCR of women visiting child health clinics	EPDS ≥12/13, MDD or minD (RDC/GSPI)	1. Counseling 2. CAU	26 24	Ind	Nondirective counseling (8 s)	Pre, post	EPDS	UK	9	+	−	
Honey, Bennett, & Morgan, 2002	Referrals by health visitor	EPDS >12	1. Psychoeducation 2. CAU	23 22	Grp	Psychoeducation, CBT, relaxation (8 s)	Pre, post, 6 min	EPDS	UK	0	−	−	
Meager & Milgrom, 1996	Community recruitment	EPDS>12+BDI>15	1. Support group 2. WL	10 10	Grp	Support+education+CBT (10 sessions)	Pre, post	EPDS, BDI, POMS	AU	40	−	−	+
Milgrom, Negri, Gemmill, McNeil, & Martin, 2005	SCR in maternal and child health centers	EPDS ≥12+MDD of minD (DSM-IV/CIDI)	1. CBT group 2. Counseling, group 3. Counseling, individual 4. CAU	46 47 66 33	Grp/Ind	9, 90 min/week, 3 s w/partner, 12 weeks	Pre, post, 12 min	EPDS BDI	AU	37	+	−	+
Misri, 2004	Referrals to hospital	HRSD ≥18+EPDS ≥20+DD (DSM-IV)	1. CBT+paroxetine 2. Paroxetine	19 16	Ind	Cognitive behavior therapy (12 s)	Pre, post	HRSD	CAN	20	+	+	+

M. H. O'Hara, Stuart, Gorman, & Wenzel, 2000	Community SCR	HDRS ≥12 +MDD (DSM-IV/ SCID)	1. IPT 2. WL	60 60	Ind	Interpersonal psychotherapy (12 s)	Pre, post	HDRS, BDI, EPDS	US	17	−	−	+
Prendergast & Austin, 2001	SCR by early childhood nurses	EPDS >12+ MDD or minD (DSM-IV)	1. CBT 2. Advice+ support	17 20	Ind	Cognitive-behavior therapy (6 s)	Pre, post, 6 min	EPDS, MADRS	AU	NR[a]	−	−	+
Tezel, 2006	SCR during postnatal care	EPDS >11, no symptoms of MDD	1. Problem-solving 2. Systematic care	30 32	Ind	Problem-solving training (6 s)	Pre, post	BDI, EPDS	Tur	0	−	−	−
Wickberg, 1996	Community SCR	EPDS ≥12+ MADRS ≥10	1. Counseling 2. CAU	20 21	Ind	Nondirective counseling (6 s)	Pre, post	EPDS, ADRS	Swe	15	−	−	+

[a]The abbreviations of the measurement instruments are not reported here; the reader is referred to the primary studies.
[b]+indicates blinding of assessors of outcome;−indicates no blinding, or blinding of assessors is not reported.
[c]In these studies, it was not reported whether there was no attrition or that only intention-to-treat analyses on the full sample were conducted.
All = adequacy of random allocation concealment to respondents; Attr = attrition; AU = Australia; BA = blinding of assessors of outcome; CAN = Canada; CAU = care-as-usual; CBT = cognitive behavior therapy; FR = France; Grp = group; Ind = individual; IPT = interpersonal psychotherapy; MDD = major depressive disorder; minD = minor depression; NR = not reported; PLA = placebo; RA = random assignment to conditions; S = sessions; SCR = systematic screening; Swe = Sweden; Ther = therapy; Tur = Turkey; UK = United Kingdom; US = United States; WL = waiting list control group.

simple approach for assessing the validity of the studies, as suggested in the Cochrane Handbook (Higgins & Green, 2005). We assessed the validity of the studies using four basic criteria: allocation to conditions is done by an independent (i.e., third) party; adequacy of random allocation concealment to respondents; blinding of assessors of outcomes; and completeness of follow-up data.

The methodological quality of the included studies was not optimal. Allocation to conditions was conducted by an independent (i.e., third) party in only four studies. Random allocation concealment to respondents was not possible or not reported in all studies. Blinding of assessors of outcomes was reported in seven studies. Attrition rates ranged from 0 to 40%.

Effects at Posttest

We calculated effect sizes (d) by subtracting (at posttest) the average score of the control group (M_e) from the average score of the experimental group (M_c) and dividing the result by the average of the standard deviations of the experimental and control groups (SD_{ec}). An effect size of 0.5 thus indicates that the mean of the experimental group is half a standard deviation larger than the mean of the control group. Effect sizes of .56 to 1.2 can be assumed to be large, effect sizes of .33 to .55 are moderate, and effect sizes of 0 to .32 are small (Lipsey & Wilson, 1993).

In the calculations of effect sizes, we used only those instruments that explicitly measure depression (see Table 1), such as the EPDS (Cox et al., 1987) and the BDI (A. T. Beck et al., 1961). If more than one depression measure was used, the mean of the effect sizes was calculated so that each study (or contrast group) had only one effect size. In three studies, more than one experimental condition was compared to a control condition (Clark, Wenzel, & Tluczek, 2003; Cooper et al., 2003; Milgrom, Negri, Gemmill, McNeil, & Martin, 2005). In these cases, the number of participants in the control condition was evenly divided over the experimental conditions so that each participant was used only once in the meta-analyses.

To calculate pooled mean effect sizes, we used the computer program Comprehensive Meta-analysis (Version 2.2.021), developed for support in meta-analysis. As we expected considerable heterogeneity, we decided to calculate mean effect sizes with the random effects model. In the random effects model, it is assumed that the included studies are drawn from ''populations'' of studies that systematically differ from each other. In this model, the effect sizes resulting from included studies differ because of the random error within studies, but also because of true variation in effect size from one study to the next.

In our analyses, we tested whether there are genuine differences underlying the results of the studies (heterogeneity), or whether the variation in findings is compatible with chance alone (homogeneity; Higgins, Thompson, Deeks, & Altman, 2003). As an indicator of homogeneity, we calculated the Q-statistic. A significant Q rejects the null hypothesis of homogeneity and indicates that the variability among the effect sizes is greater than what is likely to have resulted from subject-level sampling error alone. We also calculated the I^2 statistic, which is an indicator of heterogeneity in percentages as well. 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 (Higgins et al., 2003).

The effects of the psychological treatments of PPD could be compared to control conditions in 19 comparisons from 14 studies (Table 2). The mean effect size of all studies was 0.61 (95% CI: 0.37~0.85; $Z = 5.06$, $p < .001$), with moderate to high

heterogeneity ($Q = 51.20$, $p < .001$; $I^2 = 64.84$). The effect sizes of all studies are presented in Figure 1.

Visual inspection of Figure 1 suggested that the study by Chabrol and colleagues (2002) was an outlier because the 95% confidence interval around the effect size did not overlap with the other studies (except one). We excluded this study to examine whether this would reduce heterogeneity. The resulting effect size was 0.51 (95% CI: 0.34~0.68; $Z = 5.87$, $p < .001$) with low heterogeneity ($Q = 24.81$, n.s., $I^2 = 31.47$). Because of this large reduction in heterogeneity, we decided to remove this study from all further analyses.

Because three studies had more than one comparison and these effect sizes are not independent, we examined whether this influenced the results of our study. We conducted a new meta-analysis in which we selected the smallest effect size from each study while the other effect sizes from these three studies were removed. The resulting mean effect size was 0.56 (95% CI: 0.34~0.79; $Z = 4.91$, $p < .001$; 13 comparisons), with somewhat higher (but still moderate) heterogeneity ($Q = 22.70$, $p < .05$; $I^2 = 47.13$). This was comparable to the results of the overall analyses, indicating that these studies did not result in an overestimation of the effect size.

Publication Bias

Publication bias was tested by inspecting the funnel plot on primary outcome measures (i.e., effects on depression at posttest), and by Duval and Tweedie's (2000) trim and fill procedure, which yields an estimate of the effect size after the publication bias has been taken into account (as implemented in Comprehensive Meta-analysis, Version 2.2.021). Neither the funnel plots nor Duval and Tweedie's trim and fill procedure indicated a significant publication bias. The effect size indicating the difference in depressive symptomatology between experimental and control conditions did not change after adjustment for possible publication bias (i.e., the observed and adjusted d were the same). We also calculated Orwin's fail-safe N, which is the number of studies with an effect size of zero that should be found to reduce the mean effect size to 0.20, and this was found to be 28.

Subgroup Analyses

Because some heterogeneity was present in the analyses, we examined whether the effect sizes of specific subgroups differed from each other, with the methods for subgroup analyses as implemented in Comprehensive Meta-analysis Version 2.2.021. We used mixed effects analyses, which pooled studies within subgroups with the random effects model but tested for significant differences between subgroups with the fixed effects model. We selected a number of characteristics of the target population, the intervention, and the general design of the studies for the subgroup analyses. Criteria for selecting these characteristics were (a) the characteristic had to be a core element of the target population, the intervention, or the design of the studies; and (b) it had to result in subgroups with three or more effect sizes. This resulted in the following subgroups: recruitment method (systematic screening vs. other recruitment methods); definition of depression (a diagnosed depressive disorder vs. a high score on a self-report instrument); a cognitive behavioral intervention (yes/no); a social support intervention (yes/no); intervention format (individual or group); type of control group (care-as-usual vs. other control conditions); method of effect size computation (directly from the mean and standard deviation or indirectly through other statistics reported in the paper); and random

assignment to experimental or control group (yes/no). The results of these analyses are presented in Table 2.

The subgroup analyses indicated that studies in which a waiting list condition was used had significantly higher effect sizes than did studies in which care-as-usual or other control groups were used ($p < .05$). We also found a trend indicating that the effect sizes in the three studies in which participants were not randomly assigned to the experimental and control condition were smaller than those for the studies in which random assignment was used. The group of three studies without random assignment and the group of two studies in which another (i.e., no waiting or care-as-usual) control group was used were the only subgroups with a nonsignificant mean

Table 2

		N	*d*	95% CI	*Z*	*Q*	I^2	*p*
Overall analyses								
All studies		19	0.61	0.37∼0.85	5.06***	51.20**	64.84	
All studies, 1 outlier excluded[a]		18	0.51	0.34∼0.68	5.87***	24.81	31.47	
Subgroup analyses								
Recruitment	Systematic screening	14	0.51	0.31∼0.71	4.92***	23.58*	44.87	n.s.
	Other	4	0.53	0.13∼0.92	2.60**	1.21	0	
Diagnosis	Depressive disorder	12	0.53	0.33∼0.72	5.31***	13.38	17.79	n.s.
	Other	6	0.54	0.19∼0.90	2.98**	10.83	53.82	
CBT	Yes	8	0.36	0.15∼0.58	3.31**	2.56	0	n.s.
	No	10	0.64	0.36∼0.93	4.50***	19.63*	54.15	
Social support intervention	Yes	4	0.54	0.03∼1.05	2.07*	8.85*	66.11	n.s.
	No	14	0.54	0.37∼0.71	6.13***	14.67	11.40	
Intervention format	Group	6	0.42	0.10∼0.73	2.56*	7.48	33.16	n.s.
	Individual	9	0.64	0.41∼0.87	5.53***	11.37	29.62	
Control condition	Care-as-usual	12	0.41	0.25∼0.58	4.85***	11.70	5.99	*
	Waiting-list	4	0.96	0.63∼1.29	5.65***	2.69	0	
	Other	2	0.40	–0.17∼0.97	1.38 n.s.	1.64	38.95	
Effect size computation	Direct	10	0.46	0.20∼0.72	3.46**	18.87*	52.32	n.s.
	Indirect	8	0.59	0.37∼0.82	5.13***	4.91	0	
Attrition rate	≤20%	11	0.55	0.31∼0.80	4.47***	19.86*	49.65	n.s.
	>20%	5	0.43	0.12∼0.74	2.72**	2.51	0	
Random assignment	Yes	15	0.58	0.41∼0.75	6.82***	16.15	13.33	0
	No	3	0.20	−0.18∼0.59	1.03 n.s.	2.46	18.69	
Effects at follow-up								
All studies	6–12 months' follow-up	8	0.48	0.05∼0.90	2.20*	17.40*	59.76	
Outlier excluded[b]	6–12 months' follow-up	5	0.16	−0.11∼0.42	1.16 n.s.	2.92	0	
Comparisons to other treatments								
Other treatments	At posttest	3	−0.86	−1.45∼−0.28	−2.88**	4.49 n.s.	55.50	
Psychological vs. combined	At posttest	2	−0.17	−0.62∼0.28	−0.75 n.s.	0.32 n.s.	0	

*$p < .05$; **$p < .01$; ***$p < .001$.
[a]Chabrol et al., 2002. [b]Milgrom et al., 2005.
n.s. = not significant.

effect size (i.e., not significantly different from zero). None of the other subgroup analyses were significant.

Effects at Follow-Up

Four studies reported the effects of a psychological treatment compared to a (care-as-usual) control group at follow-up. The follow-up periods ranged from 6 to 12 months. One study reported follow-up effects after 5 years (Cooper et al., 2003). We could calculate effect sizes for eight comparisons from the four studies between intervention and control conditions at 6 to 12 months' follow-up. The overall effect size was 0.48 (95% CI: 0.05∼0.90; $Z = 2.20$, $p < .05$) with moderate to high heterogeneity (Q = 17.40, $p < .05$; $I^2 = 59.76$). Because the attrition rate was very high in one study (i.e., 57 of the 192 cases who started with the study; 29.7%; Milgrom et al., 2005) and because visual inspection of the funnel plot suggested that this may be an outlier, we excluded this study. The resulting effect size was 0.16 (95% CI: −0.11∼0.42; $Z = 1.16$, n.s.) with zero heterogeneity (Q = 2.92, n.s., $I^2 = 0$).

Other Comparisons

Psychological treatments were compared to other interventions (i.e., pharmacological treatment, systematic care, pram-walking) in three studies, with a resulting mean effect size of −0.86 (95% CI: −1.45∼−0.28; $Z = -2.88$, $p < .01$), indicating superior effects of the other treatments. However, these interventions differed strongly from each other, and as could be expected, heterogeneity was relatively high (Q = 4.49, n.s.; $I^2 = 55.50$). In two studies, the effects of psychological treatments were compared to combined psychological and pharmacological treatments ($d = -0.17$; 95% CI: −0.62∼0.28; $Z = -0.75$, n.s.; Q = 0.32, n.s.; $I^2 = 0$).

Discussion

We found evidence that psychological treatments have moderate effects on depression in women with PPD. Accordingly, psychological treatments do have significant effects in PPD, but the effect size that we found indicated somewhat smaller effects than are usually found for psychological treatments (Churchill et al., 2001; Cuijpers & Dekker, 2005). Biological treatments such as pharmacological treatment and electroconvulsive therapy also find higher effect sizes. It is not clear why the effects of psychological treatments are smaller in PPD. One possible explanation is that biological causes are more prominent in PPD. Another explanation of the smaller effect sizes of psychological treatments for PPD is that most studies in this area used care-as-usual control groups. It is known from other research that care-as-usual control groups typically result in smaller effects than those of waiting list control groups (Cuijpers, van Straten, Warmerdam, & Smits, in press). This was confirmed in our study. We could not find evidence that psychological treatments have a significant effect on PPD in the longer term (i.e., at 6 to 12 months' follow-up), but this may be due to the very small number of studies examining longer term effects.

The few studies that compared psychological treatments to other treatments indicated that the other treatments were somewhat more effective. This could be an indication that psychological treatments may not be the treatment of first choice for PPD, but this can only be a cautious hypothesis as the number of studies was too small to draw any definite conclusions.

We found indications for some heterogeneity, indicating that there are some systematic differences between the studies. Although heterogeneity was low to moderate, we conducted several subgroup analyses. Another reason to conduct subgroup analyses was that in some studies, more comparisons between psychological treatments and control conditions were made. These comparisons were not independent, and when we controlled for these comparisons, heterogeneity was larger (although still moderate).

The subgroup analyses indicated that studies with waiting list control conditions had larger effects than studies in which care-as-usual or other control groups were used. Both the subgroup of studies with waiting list control groups and the subgroup of studies with care-as-usual control groups had very low levels of heterogeneity, and this may explain a considerable part of the heterogeneity in the sample of studies. We also found a trend indicating that studies in which participants were not assigned to conditions at random had smaller effect sizes than studies that used randomized designs. Both subgroups had low levels of heterogeneity. Accordingly, random assignment may be another characteristic that is responsible for the heterogeneity. Because the number of comparisons was very small, however, this has to be interpreted very cautiously.

We included a broad range of psychological interventions for PPD, including cognitive behavior therapy, interpersonal psychotherapy, and counseling, but also social support interventions in which peer support was a core element. Despite this broad range of interventions, we did not find indications that effect sizes of the major categories (i.e., cognitive behavior therapy vs. other interventions, social support interventions vs. other interventions) differed from each other. The finding that different types of psychological interventions are about equally effective in the treatment of depression also was noted in other meta-analyses (Churchill et al., 2001; Cuijpers et al., in press; Wampold, Minami, Baskin, & Callen Tierney, 2002). In the general psychotherapy literature, it has been debated for more than three decades now whether psychotherapies are actually equally effective (Cuijpers, 1998; Luborsky, 1995; Luborsky, Singer, & Luborsky, 1975; Shadish & Sweeney, 1991), but no definite answer has been found. It is possible that most effects of psychological treatments are caused by common, nonspecific factors and not by specific techniques, but it also is possible that the effects of psychotherapy are realized by different, therapy-specific mechanisms (Butler & Strupp, 1986) and that the number of possible mediators and moderators is so large that small differences between treatments in specific groups of patients remain unnoticed.

It was not clear why the study by Chabrol et al. (2002) was an outlier in the main analyses. This study was the only one conducted in France, where, although unlikely, care-as-usual may differ from care-as-usual in other countries. This study also was one of the studies in which the participants were not assigned at random to the treatment conditions. Another difference is that in this study, urgent cases in the control group were assessed weekly. More research is needed to examine whether these differences are actually related to the effects of the intervention.

This study has several limitations. First, the number of included studies is small. This is important for the overall analyses, but even more so for the subgroup analyses. Second, hardly any studies were found that compared psychological treatments to pharmacological and other treatments. Therefore, no conclusions can be drawn about the longer term effects or the comparative effects of psychological treatments. Third, we found that the quality of several included studies was not optimal. Although it is clearly inherent in studies of psychological treatments that it

is not possible to conceal to participants to which condition they are assigned (in waiting list control conditions, it is not possible at all), several studies did not meet other major quality criteria such as assignment to conditions by an independent person and blinding of assessors.

Although there is no doubt that psychological treatments can be an effective treatment for PPD, more research is needed to examine whether psychological treatments are actually less effective in PPD than in other depressive disorders, or that the smaller effect sizes are caused by the fact that most studies used care-as-usual control conditions. Furthermore, it is important to compare the effects of other treatments, such as pharmacotherapy, to the effects of psychological treatments to examine whether psychological treatments are indeed not the treatment of first choice in PPD. It also is important to examine more thoroughly the long-term effects of the treatments, as past research does not provide clear indications of how long effects of the treatments last.

References

Appelby, L., Warner, R., Whitton, A., & Faragher, B. (1997). A controlled study of fluoxetine and cognitive-behavioral counselling in the treatment of postnatal depression. British Medical Journal, 314, 932–936.

Armstrong, K., & Edwards, H. (2004). The effectiveness of a pram-walking exercise programme in reducing depressive symptomatology for postnatal women. International Journal of Nursing Practice, 10, 177–194.

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

Beck, C.T. (1995). The effects of postpartum depression on maternal–infant interaction: A meta-analysis. Nursing Research, 44, 298–304.

Beck, C.T. (1999). Maternal depression and child behaviour problems: A meta-analysis. Journal of Advances in Nursing, 29, 623–629.

Buist, A. (1998). Childhood abuse, parenting and postpartum depression. Australian and New Zealand Journal of Psychiatry, 32, 479–487.

Butler, S.F., & Strupp, H.H. (1986). Specific and nonspecific factors in psychotherapy: A problematic paradigm for psychotherapy research. Psychotherapy: Theory, Research and Practice, 23, 30–40.

Chabrol, H., Teissedre, F., Saint-Jean, M., Teisseyre, N., Roge, B., & Mullet, E. (2002). Prevention and treatment of post-partum depression: A controlled randomized study on women at risk. Psychological Medicine, 32, 1039–1047.

Chen, C.H., Tseng, Y.F., Chou, F.H., & Wang, S.Y. (2000). Effects of support group intervention in postnatally distressed women: A controlled study in Taiwan. Journal of Psychosomatic Research, 49, 395–399.

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

Clark, R., Wenzel, A., & Tluczek, A. (2003). Psychotherapy for postpartum depression: A preliminary report. American Journal of Orthopsychiatry, 73, 441–454.

Cooper, P.J., & Murray, L. (1998). Fortnightly review: Postnatal depression. British Medical Journal, 316, 1884–1886.

Cooper, P.J., Murray, L., Wilson, A., & Romaniuk, H. (2003). Controlled trial of the short- and long-term effect of psychological treatment of post-partum depression: I. Impact on maternal mood. British Journal of Psychiatry, 182, 412–419.

Cox, A.D., Puckering, C., Pound, A., & Mills, M. (1987). The impact of maternal depression in young children. Journal of Child Psychology and Psychiatry, 28, 917–928.

Cox, J.L., Holden, J.M., & Sagovsky, R. (1987). Detection of postnatal depression: Development of the 10-item Edinburgh Postnatal Depression Scale. British Journal of Psychiatry, 150, 782–786.

Cuijpers, P. (1998). Minimising interventions in the treatment and prevention of depression: Taking the consequences of the "Dodo bird verdict." Journal of Mental Health, 7, 355–365.

Cuijpers, P., van Straten, A., & Smit, F. (2006). Psychological treatment of late-life depression: A meta-analysis of randomized controlled trials. International Journal of Geriatric Psychiatry, 21, 1139–1149.

Cuijpers, P., van Straten, A., & Warmerdam, L. (2007a). Behavioral treatment of depression: A meta-analysis of activity scheduling. Clinical Psychology Review, 27, 318–326.

Cuijpers, P., van Straten, A., & Warmerdam, L. (2007b). Problem solving therapies for depression: A meta-analysis. European Psychiatry, 22, 9–15.

Cuijpers, P., van Straten, A., Warmerdam, L., & Smits, N. (in press). Characteristics of effective psychological treatments of depression: A meta-regression analysis. Psychotherapy Research.

Cuijpers, P., & Dekker, J. (2005). Psychologische behandeling van depressie; een systematisch overzicht van meta-analysen [Psychological treatment of depression: A systematic review of meta-analyses]. Nederlands Tijdschrift voor Geneeskunde [Dutch Journal of Medicine], 149, 1892–1897.

Cummings, E.M., & Davies, P.T. (1994). Maternal depression and child development. Journal of Child Psychology and Psychiatry, 35, 73–112.

Dennis, C.L. (2003). The effect of peer support on postpartum depression: A pilot randomized controlled trial. Canadian Journal of Psychiatry, 48, 115–124.

Dennis, C.L. (2004). Treatment of postpartum depression, Part 2: A critical review of nonbiological interventions. Journal of Clinical Psychiatry, 65, 1252–1265.

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometrics, 56, 455–463.

Dennis, C.L., & Creedy, D. (2004). Psychosocial and psychological interventions for preventing postpartum depression. Cochrane Database of Systematic Reviews 2004, Issue 4. Art. No. CD001134. DOI: 10.1002/14651858.CD001134.pub2.

Elliott, S.A. (2000). Report on the Satra Bruk workshop on classification of postnatal mental disorders. Archives of Women's Mental Health, 3, 27–33.

Fleming, A.S., Klein, E., & Corter, C. (1992). The effects of a social support group on depression, maternal attitudes and behavior in new mothers. Journal of Child Psychology and Psychiatry, 33, 685–698.

Gjerdingen, D. (2003). The effectiveness of various postpartum depression treatments and the impact of antidepressant drugs on nursing infants. Journal of the American Board of Family Practice, 16, 372–382.

Goodman, S.H., & Gotlib, I.H. (1999). Risk for psychopathology in the children of depressed mothers: A developmental model for understanding mechanisms of transmission. Psychology Review, 106, 458–490.

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

Higgins, J.P.T., & Green, S. (2005). Cochrane handbook for systematic reviews of interventions 4.2.5 [updated May 2005]. In The Cochrane Library, Issue 3. Chichester, UK: Wiley.

Hipwell, A.E., Goossens, F.A., Melhuish, E.C., & Kumar, R. (2000). Severe maternal psychopathology and infant–mother attachment. Developmental Psychopathology, 12, 157–175.

Holden, J.M. (1991). Postnatal depression: Its nature, effects, and identification using the Edinburgh Postnatal Depression Scale. Birth, 18, 211–221.

Holden, J.M., Sagovsky, R., & Cox, J.L. (1989). Counselling in a general practice setting: Controlled study of health visitor intervention in treatment of postnatal depression. British Medical Journal, 298, 223–226.

Honey, K.L., Bennett, P., & Morgan, M. (2002). A brief psycho-educational group intervention for postnatal depression. British Journal of Clinical Psychology, 41, 405–409.

Lipsey, M.W., & Wilson, D.B. (1993). The efficacy of psychological, educational and behavioral treatment. American Psychologist, 48, 1181–1209.

Luborsky, L. (1995). Are common factors across different psychotherapies the main explanation for the Dodo bird verdict that "everybody has won so all shall have prizes?" Clinical Psychology: Science and Practice, 2, 106–109.

Luborsky, L., Singer, B., & Luborsky, L. (1975). Comparative studies of psychotherapies: Is it true that "everybody has won and all must have prizes?" Archives of General Psychiatry, 32, 995–1008.

Lumley, J., Austin, M.P., & Mitchell, C. (2004). Intervening to reduce depression after birth: A systematic review of the randomized trials. International Journal of Technology Assessment in Health Care, 20, 128–144.

Meager, I., & Milgrom, J. (1996). Group treatment for postpartum depression: A pilot study. Australian and New Zealand Journal of Psychiatry, 30, 852–860.

Milgrom, J., Negri, L.M., Gemmill, A.W., McNeil, M., & Martin, P.R. (2005). A randomized controlled trial of psychological interventions for postnatal depression. British Journal of Psychology, 44, 529–542.

Misri, S., Reebye, P., Corral, M., & Milis, L. (2004). The use of paroxetine and cognitive-behavioral therapy in postpartum depression and anxiety: A randomized controlled trial. Journal of Clinical Psychiatry, 65, 1236–1241.

Murray, L. (1992). The impact of postnatal depression on infant development. Journal of Child Psychology and Psychiatry, 33, 543–561.

Murray, L., & Cooper, P.J. (1997). Postpartum depression and child development. New York: Guilford Press.

Murray, L., & Cooper, P.J. (2004). The impact of postpartum depression on child development. In I. Goodyer (Ed.), Aetiological mechanisms in developmental psychopathology. Oxford, England: Oxford University Press.

O'Hara, M.H. (1997). The nature of postpartum depressive disorders. In L. Murray & P. J. Cooper (Eds.), Postpartum depression and child development (pp. 3–31). New York: Guilford Press.

O'Hara, M.W., & Swain, A.M. (1996). Rates and risk of postpartum depression: A meta-analysis. International Review of Psychiatry, 8, 37–54.

O'Hara, M.W., Stuart, S., Gorman, L.L., & Wenzel, A. (2000). Efficacy of interpersonal psychotherapy for postpartum depression. Archives of General Psychiatry, 57, 1039–1045.

Orvaschel, H., Walsh-Allis, G., & Ye, W.J. (1988). Psychopathology in children of parents with recurrent depression. Journal of Abnormal Child Psychology, 16, 17–28.

Prendergast, J., & Austin, M.P. (2001). Early childhood nurse-delivered cognitive behavioural counselling for post-natal depression. Australasian Psychiatry, 9, 255–259.

Richards, J.P. (1990). Postnatal depression: A review of recent literature. British Journal of General Practice, 40, 472–476.

Shadish, W.R., & Sweeney, R.B. (1991). Mediators and moderators in meta-analysis: There's a reason we don't let dodo birds tell us which psychotherapies should have prizes. Journal of Consulting and Clinical Psychology, 59, 883–893.

Sit, D., Rothschild, A.J., & Wisner, K.L. (2006). A review of postpartum psychosis. Journal of Women's Health, 15, 352–368.

Stocky, A., & Lynch, J. (2000). Acute psychiatric disturbance in pregnancy and the puerperium. Baillieres Best Practice Research in Clinical Obstetric Gynaecology, 14, 73–87.

Tezel. A., & Gözüm, S. (2006). Comparison of effects of nursing care to problem solving training on levels of depressive symptoms in post partum women. Patient Education and Counseling, 63, 64–73.

Wampold, B.E., Minami, T., Baskin, T.W., & Callen Tierney, S. (2002). A meta-(re)analysis of the effects of cognitive therapy versus "other therapies" for depression. Journal of Affective Disorders, 68, 159–165.

Whiffen, V.E., & Gotlib, I.H. (1989). Infants of postpartum depressed mothers: Temperament and cognitive status. Journal of Abnormal Psychology, 98, 274–279.

Wickberg, B., & Hwang, C.P. (1996). Counselling of postnatal depression: A controlled study on a population based Swedish sample. Journal of Affective Disorders, 39, 209–216.

Wisner, K.L., Chambers, C., & Sit, D.K.Y. (2006). Postpartum depression: A major public health problem. Journal of the American Medical Association, 296, 2616–2618.