Authors' objectives
To evaluate treatments of postnatal depression.

Searching
MEDLINE, PsycLIT, Sociofile, CINAHL and COPAC databases, and published books held by the British Library, were searched from 1964 to 2000 for publications in the English language. The keywords used were 'postnatal depression', 'postpartum depression', 'puerperal depression' and 'treatment'. In addition, the references in retrieved publications were examined, and Marce Society conference proceedings and abstracts were searched.

Study selection
Study designs of evaluations included in the review
There were no restrictions on the study designs eligible for inclusion in the review. The included studies were RCTs, quasi-controlled studies, cohort studies, open studies, descriptive reports, case studies and case series; one study was of an unknown design.

Specific interventions included in the review
Studies that clearly defined their main purpose as treatment were eligible. The included studies evaluated the following treatments:

pharmacological treatments, i.e. fluoxetine, sertraline, and S-adensylmethionine;

combined psychological and psychodynamic approaches, i.e. interactional/problem resolution, cognitive-behaviour therapy with and without counselling, group psychotherapy, listening visits, psychodynamic psychotherapy or interactional guidance, interpersonal psychotherapy, psychotherapy on mother-infant interaction, and counselling;

combined pharmacological and psychological approaches, i.e. fluoxetine plus cognitive-behavioural counselling and specialist day hospital;

social support and relaxation, i.e. social support group, group by mail, exercise and relaxation, massage, relaxation, and support group with and without partner;

hormonal, i.e. oestradiol skin patches and sublingual oestradiol;

other approaches, i.e., bright light therapy; and

routine primary care (used as the control).

Participants included in the review
Studies of women with broadly defined postnatal depression were eligible regardless of the cause or time of onset of postnatal depression. Women with postpartum blues and puerperal psychosis were excluded.

Women were diagnosed with postnatal depression using the following criteria (where specified): American Psychiatric Association (APA) DSM-III-R and DSM-IV for major depression; Edinburgh Postnatal Depression Scale (EPDS) with and without further enquiry; Research Diagnostic Criteria (RDC) major or minor depression; Beck Depression Inventory (BDI); a score greater than 16 on BDI plus dysthymia on Diagnostic Inventory Schedule; a score greater than 35 on Current Experience Scale plus either greater than 21 on Multiple Affect Adjective Checklist or greater than 13 on EPDS; Hospital Anxiety and Depression Scale; and ICD-10 depression.
Some of the included studies also enrolled women with postnatal depression (criteria not specified), families with a child aged less than 30 months who had consulted a child guidance clinic, and women with postnatal depression plus borderline personality disorder.

Outcomes assessed in the review
The inclusion criteria were not defined in terms of the outcome. The included studies used at least 29 different instruments to assess the outcomes. The randomised controlled trials (RCTs) used the following criteria: Kellner Symptom Questionnaire (KSQ); SPI; EPDS; RDC diagnosis; BDI; Profile of Mood Status (POMS); Dyadic Adjustment Scale (DAS); Coopersmith Self-Esteem Inventory (CSEI); Social Provision Scale; Montgomery-Asberg Depression Rating Scale; APA DSM-III-R; depression section of the Structured Clinical Interview; Hamilton Depression Rating Scale (HAM-D/HDRS); Social Adjustment Scale; Inventory to Diagnose Depression; Postpartum Adjustment Questionnaire; Revised Clinical Interview Schedule; Behaviour Observation Scale; State Anxiety Instrument for Children; Mini International Neuropsychiatric Instrument; General Health Questionnaire; PBI; Parenting Stress Index (PSI); and Schedule for Affective Disorders and Schizophrenia (SADS-C). Urinary and salivary cortisol and pulse rate were also assessed.

How were decisions on the relevance of primary studies made?
The authors do not state how the papers were selected for the review, or how many of the reviewers performed the selection.

Assessment of study quality
Validity was assessed using criteria based on the criteria devised by Sackett et al.(see Other Publications of Related Interest). The criteria related to the following elements: the methods used to select the patients; the inclusion and exclusion criteria for the patients; power sample calculation; adequacy of sample size; random assignment of treatment; the method of randomisation; reporting of confounding factors; equal treatment of intervention groups; the use of a control group; baseline comparability of the groups; intervention clearly described and replicable; the degree of blinding; outcome measures validated for use in postnatal depression; clear presentation of the results; description of the drop-outs; analysis conducted on an intention to treat basis; the duration and completeness of follow-up; and the independence of the funding source or no specified funding source. Two authors independently assessed study validity according to the defined criteria. Any disagreements were discussed until agreement was reached. Clarification was sought from the original authors where possible.

Data extraction
The authors do not state how the data were extracted for the review, or how many of the reviewers performed the data extraction.

The tabulated data included the following information: author and year of publication; country where the study was conducted; sample size; study design; inclusion criteria; assessment criteria; and treatment.

Methods of synthesis
How were the studies combined?
The studies were grouped by intervention type and a narrative synthesis was undertaken.

How were differences between studies investigated?
Differences between the studies with respect to validity were discussed in the text.

Results of the review
Twenty-nine studies with at least 926 women were included. Of these, there were 9 RCTs (476 women), one quasi-controlled study (142 women) and 2 cohort studies (130 women). The other studies were either case studies or descriptive reports.
The methodological flaws included: lack of placebo control; extremely small sample size; high drop-out rates; no criteria or inadequate criteria used to diagnose postpartum depression; short follow-up periods; use of assessments instruments that were not validated for use in this population; lack of control for confounding variables and co-interventions; no assessment of compliance with treatment; selection bias; selective quotation of patient's statements in qualitative research; method of randomisation not described; inadequate methods of randomisation; lack of standardisation of interventions in multicentre trials; baseline differences between the treatment groups; no description of the timing of assessments with respect to the birth or the treatments; no report of the number of women declining the intervention; and the use of multiple assessments in control and intervention groups.

Only results from the RCTs are reported in this abstract.

Pharmacological treatments: there was one two-phase trial involving an RCT with 60 women (phase 1) and an open trial with 40 women (phase 2); one case study; and one case series. The RCT found that S-adenosylmethionine versus placebo decreased mean KSQ scores at day 10 and day 30 (the values and significance levels were not reported in the review). Methodological problems included the inclusion of women with maternity blues and postnatal depression; a short follow-up period of 30 days; measurement of outcome using a questionnaire that had not been proven to be reliable and valid during the assessment period (2 days postpartum); lack of description of the randomisation process; and baseline comparison of treatment groups on only a few socio-demographic factors.

Psychological and psychodynamic approaches: there were 5 RCTs with 401 women, one cohort study with 70 women, 8 case studies or series descriptive studies, and one study of unspecified design. The 4 RCTs found lower rates of depression for the active treatment versus control, but the studies were methodologically flawed. For example, the intervention and control intervention were carried out by the same health professionals; randomisation was not adequately described or was inadequate; some women also received antidepressants; the Hawthorne effect may have affected outcomes; selective quotes; inadequate description of intervention; high withdrawal rates; assessors not blinded; and selection bias.

One RCT (26 women) found that 8 weekly, 1-hour counselling visits by a health visitor versus usual care reduced depression rates: 69% were no longer depressed in the counselling group versus 38% in the control group. One RCT (41 women) found that 6 non-directive counselling sessions versus routine care reduced depression rates at 6 weeks: 12 out of 15 (80%) patients were no longer depressed in the counselling group versus 4 out of 16 (25%) in the routine care group. One RCT (120 women) found that interpersonal psychotherapy versus waiting-list control significantly increased the recovery rates and decreased scores on the BDI and HDRS. One RCT (a pilot study of 20 women) found that cognitive-behavioural programme versus waiting-list control significantly improved mood (EPDS, BDI, POMS), but found no significant difference between treatments in terms of DAS, CSEI or PSI. One RCT (194 women) compared non-directive counselling, cognitive behavioural therapy, dynamic psychotherapy and routine primary care, and found no significant difference between the treatments at 9 months.

Combined pharmacological and psychological approaches: there was one RCT with 87 women and one cohort study with 60 women. The RCT compared 4 treatments: fluoxetine or placebo plus either one or 6 sessions of counselling. It found that 6 sessions of counselling were better than one session in reducing psychiatric morbidity, and that women receiving fluoxetine had significantly reduced psychiatric morbidity at 12 weeks in comparison with women receiving placebo. Generalisability was limited by the high percentage of eligible women who refused to participate (101 out of 188, i.e. 54%) and high drop-out rates (26 out of 87, i.e. 30%). The study assessed outcomes using the HDRS, which has not been validated for use with puerperal women.

Social support and relaxation: there were 2 RCTs with 61 women, one quasi-controlled study with 142 women, one open study with 13 women, and 5 descriptive studies. One RCT (32 adolescent mothers) found that massage sessions versus relaxation sessions reduced anxiety levels and stress hormone levels at 10 days. However, the timing of the assessments was unclear and the interventions were not described in full. One RCT (29 women) found that a support group that included partners versus no partners significantly decreased depression rates after 7 sessions. The methodological flaws included a very small sample size, a highly selected group, the potential for interviewer bias, and short-term follow-up (1 month).

Hormonal: there was one RCT with an unclear number of women enrolled plus a preliminary report of the same study, and one case study. The RCT found that, compared with placebo, 3 months’ transdermal oestradiol followed by 3
months with additional dydrogesterone for 12 days/month increased the rate of improvement (EPDS, SADS-C). The initial drop-out rates were high (10 of the first 35 women enrolled dropped out). The women were recruited from a variety of sources.

Bright light therapy: there was one case study with 2 women.

**Authors' conclusions**

The methodological limitations of the identified studies mean that the efficacy of these treatment approaches has not been clearly established. In addition, there is very little good evidence available on which to make policy or practice recommendations. Further research is required.

**CRD commentary**

The aims of the review were stated, and the inclusion criteria were broadly defined in terms of the participants and intervention. The inclusion criteria were not defined in terms of the outcomes or study design. Several relevant sources were searched, but restricting the search to English language publications may have resulted in the omission of other relevant studies. In addition, the lack of an attempt to locate unpublished material may have resulted in publication bias.

The methods used to select the studies were not described. Validity was assessed formally using defined criteria; the methods used for the assessment were described. Methodological flaws were comprehensively discussed in the text of the review. Relevant information on the individual studies was tabulated, but overlapping populations were not highlighted in the table and the methods used to extract the data were not reported. A narrative view was appropriate in view of the heterogeneity among studies with respect to the study design, intervention and assessment of outcomes. The studies were listed and described in no apparent order, rather than being adequately summarised within each intervention category. In addition, attention was not drawn to the higher quality of evidence from RCTs. The results (i.e. values) and levels of statistical significance were not reported consistently.

The evidence presented supports the authors' conclusions.

**Implications of the review for practice and research**

Practice: The authors state that there is very little good evidence available on which to make policy or practice recommendations.

Research: The authors state that further research is required to address the management of postnatal depression.

**Bibliographic details**


**Other publications of related interest**

Record Status
This is a critical abstract of a systematic review that meets the criteria for inclusion on DARE. Each critical abstract contains a brief summary of the review methods, results and conclusions followed by a detailed critical assessment on the reliability of the review and the conclusions drawn.