MATERNAL MENTAL HEALTH & CHILD HEALTH AND DEVELOPMENT

Literature review of risk factors and interventions on Postpartum Depression

DEPARTMENT OF MENTAL HEALTH AND SUBSTANCE ABUSE

World Health Organization
2008
CHAPTER 2: DETECTION, PREVENTION AND TREATMENT OF POSTPARTUM DEPRESSION

Cindy-Lee Dennis RN PhD

Faculty of Nursing University of Toronto
©University Health Network Women’s Health Program 2003

Citation:
This chapter should be cited as:

Contact:
For further information regarding this chapter please contact:
Cindy-Lee Dennis RN PhD at cindylee.dennis@utoronto.ca
# CHAPTER 2: DETECTION, PREVENTION, AND TREATMENT OF POSTPARTUM DEPRESSION

## Table of Contents

**CHAPTER SUMMARY**

**LIST OF TABLES**

*Introduction*

*Methods*

Search Strategy

Inclusion/Exclusion Criteria

Data Abstraction and Critical Appraisal

Methodology for Synthesis

*Section I: Detection of Postpartum Depression*

Principles of Screening

Criteria for a Screening Program

Measures Used in the Detection of Postpartum Depression

Comparisons between Screening Instruments

Antenatal Screening

Screening in the Immediate Postpartum

Implications for Practice, Policy, and Research

*Section II: Prevention of Postpartum Depression*

Pharmacological Interventions

Psychological Interventions

Psychosocial Interventions

Quality Improvement Interventions

Hormonal Interventions

Other Interventions

Implications for Practice, Policy, and Research

*Section III: Treatment of Postpartum Depression*

Pharmacological Interventions

Psychological Interventions

Psychosocial Interventions

Hormonal Interventions

Other Interventions

Implications for Policy, Practice, and Research

*References*
CHAPTER SUMMARY

Introduction

Childbirth represents for women a time of great vulnerability to become mentally unwell, with postpartum mood disorders representing the most frequent form of maternal morbidity following delivery. While postpartum depression is a major health issue for many women from diverse cultures, this affective condition often remains undiagnosed resulting in limited management. The objective of this chapter is to critically review the literature to determine the current state of scientific knowledge related to the detection, prevention, and treatment of postpartum depression.

Methods

Databases relating to the medical, psychological and social science literature were searched using specific inclusion criteria and search terms to identify studies, which examined screening procedures and/or the effect of various preventive and treatment interventions on depressive symptomatology among expectant and new mothers. Randomized controlled trials, meta-analyses, and several studies with diverse designs were identified and critically appraised in order to synthesize the current findings. The search resulted in the identification of numerous postpartum depression detection studies and over 58 trials evaluating preventive and treatment intervention. The criteria used to evaluate the interventions outlined in this chapter were based on the standardized methodology developed by the Canadian Task Force on Preventive Health Care.

Key Findings

Today, both general and postpartum-specific depression instruments have been utilized to measure depressive symptomatology. By far the most widely used instrument in postpartum depression studies and for population-based screening is the Edinburgh Postnatal Depression Scale (EPDS), a 10-item self-report scale specifically designed to screen for postpartum depression in community samples. While this measure has been validated among diverse cultures resulting in varying sensitivity and specificity values, diversity and inconsistency in assessment procedures have hampered the meaningful comparison of studies and compromised the development of a cumulative body of knowledge. Although these psychometric limitations are not unique to the EPDS, the methodological explanations justify only some of the discrepancies found between the EPDS translation and validation investigations. Significant differences in the proportion of high EPDS scores across different cultural contexts were noted in an international multi-site study suggesting that cultural factors merit more attention. In addition, further research is required to determine if indeed the EPDS is the most appropriate screening instrument as new measures are being developed based on qualitative investigations.

While determining the most appropriate instrument to detect postpartum depression is challenging, immense efforts have also been undertaken to identify pregnant women who are at-risk of developing postpartum depression such that secondary preventive interventions may be implemented. A recent
systematic review of 16 antenatal screening studies, where sufficient data was available to calculate specific screening properties, was conducted. No screening instrument met the researchers’ outlined criteria for routine application in the antenatal period and the unacceptably low positive predictive values in the included studies make it difficult to recommend the use of screening tools in routine antenatal care. It is noteworthy that the predictive power of maternal mood in the immediate postpartum period (e.g., first 2 weeks postpartum) in the development of postpartum depression has consistently been reported and warrants further investigation.

The overarching question – whether screening and subsequent management is superior to management based on usual means of identification as ‘high-risk’— is controversial. It is equivocal whether further support beyond identification improves management adherence and clinical outcomes. The trade-offs between benefits and harms are an important component in the decision of whether to screen or not. Currently, there is limited information about the harms of screening and despite a wealth of studies concerning the prevalence of postpartum depression and screening accuracy, key elements of the evidence base for screening remains insufficiently developed. As such, a strong recommendation to implement screening procedures cannot be justified until further research has been completed.

The long-term consequences of postpartum depression suggest preventive approaches are warranted. Manipulation of a risk factor may improve the associated likelihood of developing postpartum depression through many different ways. The most obvious is to decrease the amount of exposure to a given risk factor or, alternatively, reduce the strength or mechanism of the relationship between the risk factor and postpartum depression. However, translating risk factor research into predictive screening protocols and preventive interventions has met with limited success, as complex interactions of biopsychosocial risk factors with individual variations need to be contemplated. Numerous studies have been examined in this review with the diverse aetiology of postpartum depression reflected in the broad range of approaches considered. Although theoretical justifications for many of these approaches have been presented, methodological limitations render intervention efficacy equivocal with scant evidence available to guide practice or policy recommendations. Despite the recent upsurge of interest in this area, many questions remain unanswered resulting in a myriad of research implications. Similarly, definite conclusions cannot be reached about the relative effectiveness of treatment approaches due to the lack of well-designed investigations. Randomized controlled trials with large and representative samples are needed to compare different treatment modalities, examine the effectiveness of individual treatment components, and determine which treatments are most useful for women with different risk factors or clinical presentations of postpartum depression.

Implications

Even though diverse measures have been created to detect depressive symptomatology, the development of a postpartum depression screening program requires careful consideration. Evidence-based decisions need
to be made regarding: (1) the most accurate screening test that is culturally sensitive, quick to administer, easy to interpret, and readily incorporated into practice; and (2) health care system issues such as cost-effectiveness, potential harm, and policies for referral. Auspiciously, research suggests postpartum depression is amenable to preventive and treatment interventions, thus providing a rationale for the development of a screening program. However, limited research has been conducted demonstrating screening improves clinical outcomes. Furthermore, few well-designed randomized controlled trials have been conducted to effectively guide practice and policy recommendations and further research is warranted if evidence-based programs are to be implemented. As there is no single etiological pathway by which women develop postpartum depression, it is improbable that a single preventive/treatment modality will be effective for all women. A multifactorial approach, which combines the contributions of the psychological, psychosocial, and biological factors, is likely to be most beneficial as it recognizes various etiological factors and individual variations.
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-1. Validation and/or translation of the Edinburgh Postnatal Depression Scale</td>
<td>88</td>
</tr>
<tr>
<td>2-2. Antenatal screening studies</td>
<td>101</td>
</tr>
<tr>
<td>2-3. Postpartum depression preventive studies</td>
<td>126</td>
</tr>
<tr>
<td>2-4. Summary quality of evidence and practice recommendations for preventive interventions</td>
<td>139</td>
</tr>
<tr>
<td>2-5. References related to commercially available antidepressant use during pregnancy or breastfeeding</td>
<td>144</td>
</tr>
<tr>
<td>2-6. Postpartum depression treatment studies</td>
<td>157</td>
</tr>
<tr>
<td>2-7. Summary recommendations for treatment interventions</td>
<td>169</td>
</tr>
</tbody>
</table>
Introduction

Childbirth represents for women a time of great vulnerability to become mentally unwell, with postpartum mood disorders representing the most frequent form of maternal morbidity following delivery (Stocky & Lynch, 2000). These affective disorders range in severity from the early maternity blues to postpartum psychosis, a serious state affecting less than 1% of mothers (Evins & Theofrastous, 1997). Along this spectrum is postpartum depression, a condition often exhibiting the disabling symptoms of dysphoria, emotional lability, insomnia, confusion, anxiety, guilt, and suicidal ideation. Frequently exacerbating these indicators are low self-esteem, inability to cope, feelings of incompetence, and loneliness (Beck, 1992; Mills, Finchilescu, & Lea, 1995; Ritter, Hobfoll, Lavin, Cameron, & Hulsizer, 2000). While postpartum depression is a major health issue for many women from diverse cultures (Affonso, De, Horowitz, & Mayberry, 2000) and has well documented public health consequences, this affective condition often remains undiagnosed resulting in limited management. The objective of this chapter is to critically review the literature to determine the current state of scientific knowledge related to the detection, prevention, and treatment of postpartum depression.

Methods

Search Strategy

Databases searched for this specific review included Medline, PubMed, CINAHL, PsycINFO, EMBASE, ProQuest, the Cochrane Library, and the WHO Reproductive Health Library from 1966 to present. As part of the quality assessment process and to measure the capture rate of relevant references, tables of contents for key journals were hand searched for the past 2 years, reference lists of included studies and relevant reviews were examined, and key postpartum depression researchers from the U.S. and Australia were contacted via email. The initial search was based on the identification of titles containing appropriate combination of keywords (Appendix D). Finally, all abstracts related to the combination of the keywords postpartum/postnatal depression and randomized controlled trials were reviewed to ensure all potentially significant interventions were reviewed. In total, approximately 500 abstracts were examined for inclusion suitability.

Inclusion/Exclusion Criteria

While there is considerable postpartum depression research in progress, the literature review involved systematically searching for published peer-reviewed articles available in English from 1990 to 2002, although select earlier studies were included based on methodological quality and/or the absence of more recent work. Research studies that focused on postpartum depression (i.e., inception of depression within the first year postpartum) were reviewed; other childbirth-related mental health disorders (i.e., pregnancy or

Cindy-Lee Dennis, PhD 77
postpartum anxiety, maternity blues, puerperal psychosis) were not appraised. Finally, research studies evaluating preventive interventions must have incorporated a postpartum depression outcome assessment beyond the first week postpartum to be included.

Data Abstraction and Critical Appraisal

In the initial stage of the search process, peer-reviewed publications were identified and potentially relevant abstracts, which met the predetermined eligibility criteria, were subsequently extracted for further examination. Research articles were then selected and assessed in a more rigorous manner to determine inclusion suitability. These articles were either included or excluded and further sub-grouped. The critical review process consisted of assessing the disorder definition (i.e., diagnostic/screening criteria used), population sampled (i.e., inclusion/exclusion criteria, recruitment process, sample size, participant characteristics), research design (i.e., control for potential bias, method and timing of assessment, statistical analysis, outcome measures, length of follow-up), level and quality of evidence, and critical analysis of variations between findings of pertinent studies.

Methodology for Synthesis

Interventions included were evaluated according to the published criteria used by the Canadian Task Force on Preventive Health Care (CTFPHC, 2003). See Overall Methodological Framework. Based on this methodology, the following chapter is comprised of three distinct sections: (1) detecting postpartum depression, (2) preventive interventions, and (3) treatment options.

Section I: Detection of Postpartum Depression

Postpartum depression is a serious mood disorder affecting many women from diverse cultures. Despite the long-standing recognition of this condition, it represents a largely undetected form of maternal morbidity. The reasons for this are twofold. First, women are often reluctant to seek professional help (Small, Brown, Lumley, & Astbury, 1994). Even though mothers have various interactions with health professionals in the postpartum period, they are frequently unwilling to disclose emotional problems, particularly depression (Brown & Lumley, 2000). One explanation for this hesitancy may be the popular myth that equates motherhood with happiness and the idealisation of the “good mother” where feelings of joy are emphasised while unhappiness is minimised. In addition, many women have difficulty understanding the problems they are experiencing, often assuming these struggles are a normal part of motherhood. For these women, the onset of symptoms may be attributed to causes other than depression, such as fatigue or relationship difficulties (Small et al., 1994; Whitton, Appleby, & Warner, 1996). Conversely, some women recognize the symptoms as depression but fear the potential help-seeking consequences such as being labelled mentally ill or an unfit mother. Even after women have made the decision to seek professional help they frequently report
feelings of embarrassment, disappointment, and frustration as health professionals may minimize their symptoms or portray their experiences as normal (Beck, 1993). It should also be recognized that not knowing where to obtain assistance is another important help-seeking barrier (McIntosh, 1993). Finally, family members may discourage women from seeking help, as in some cultures it is unacceptable to admit to depressive symptoms or discuss such difficulties external to the family context (Matthey, Barnett, & Elliott, 1997; Okano, Nagata, Hasegawa, Nomura, & Kumar, 1998).

Health professionals may also contribute to the under-diagnosing of postpartum depression. Many health professionals have limited training in the assessment or management of postpartum depression. As such, they often do not recognize the presenting symptoms as indicating depression or they may feel uncertain about how to effectively assist and are therefore reluctant to raise such issues. However, research suggests that screening may significantly assist health professionals in their ability to detect postpartum depression. In a US study, 391 mothers were assigned to either a postpartum screening group, where the Edinburgh Postnatal Depression Scale (EPDS) was administered, or a control group, which consisted of spontaneous detection via routine clinical examination (Evins, Theofrastous, & Galvin, 2000). As expected, the incidence of depressive symptomatology detection was significantly higher in the screening group than in the spontaneous detection group (35.4% vs. 6.3% respectively; \( p < 0.001 \)). Similar findings were found in another US study where women who completed the EPDS were significantly more likely to be identified with postpartum depression symptomatology than those in the routine examination group: 11 of 37 women (30%) versus 0 of 35 women (\( p < 0.001 \)) (Fergerson, Jamieson, & Lindsay, 2002); other researchers have found comparable results (Georgiopoulous, Bryan, Wollan, & Yawn, 2001; Hearn et al., 1998). Complementing these empirical findings are interviews with physicians and midwives participating in postpartum depression screening programs, which indicate administering the EPDS not only increases awareness but also promotes appropriate referrals (Schaper, Rooney, Kay, & Silva, 1994). These preceding results suggest that the incorporation of a screening tool into clinical practice can improve health professional responsiveness and may be an effective adjunct to postpartum assessments.

**Principles of Screening**

Before symptoms are readily identifiable by health professionals, serious diseases or conditions may be present in affected individuals without their knowledge. Screening is the most widely used method for early detection and is defined as the positive identification of unrecognized disease or defect through the application of tests, examinations, or other procedures that can be rapidly applied (Shah, 1998). However, it should be noted that a positive screening result does not always equate to possessing the targeted condition, as screening procedures are not diagnostic. Therefore, to ensure clinical utility screening tests are evaluated in terms of their validity, which is established through accurate diagnostic methods and expressed in terms of sensitivity and specificity. Sensitivity refers to the ability of the test to identify correctly individuals who...
truly have the condition while specificity refers to the ability of the test to identify correctly individuals who do not have the condition. Logically, tests with low sensitivity and specificity are considered ineffective screening tools. However, a screening test is never 100% sensitive and specific, as high sensitivity is gained at the expense of specificity and vice versa. Validity is also determined through a positive predictive value, which is the proportion of individuals screened positive by a test that actually have the condition (i.e., the proportion of true positives in all test positives). It is notable that positive predictive values tend to be higher when the condition is more prevalent in the target population. Together, these main psychometric characteristics assist health professionals in determining the clinical utility of a screening tool.

Criteria for a Screening Program

As new screening tests arise, pressures to adopt and institutionalize screening programs emerge. However, according to diverse experts (Cadman, Chambers, Feldman, & Sackett, 1984; Sackett, 1987), including the Joint World Health Organization (WHO)/International Agency for Research on Cancer (IARC) screening program implementation criteria, screening procedures are only justifiable if the following standards are met:

1. Disease Issues
   1.1. Conditions/diseases for which screening is used should be important health problems. If there is an extremely low incidence of a condition, the cost and effort of screening may be prohibitive. Understanding the incidence and prevalence of the condition in a population is necessary before embarking on any large-scale screening program.
   1.2. The progression of the condition should be understood; if controlled studies have demonstrated that the natural history of the condition is not favourably altered by early detection and management then screening should not be instituted.
   1.3. Effective treatment for individuals with the conditions should be available.

2. Screening Test Issues
   2.1. Screening tests should have good sensitivity, specificity, and predictive value.
   2.2. The screening procedure should be safe, convenient, and acceptable to the target population.
   2.3. Screening tests should be cost-effective, easy to interpret, and readily incorporated into practice.
   2.4. Screening tests should be accessible to the target population.

3. Health System Issues
   3.1. A clearly defined population should be targeted.
   3.2. Comparing the costs and efficiency of various screening procedures for a condition is necessary for achieving maximum benefits at minimum cost.
3.3. An analysis of harms and benefits should be conducted. (i.e., overall long-term benefits should be greater than long-term detriment).

3.4. Strategies should be in place to ensure that the screening program will reach those who will benefit the most from the program.

3.5. Policies should stipulate what action should be taken for borderline results in order to avoid over-identifying the condition.

3.6. Standard policies for referral and preventive/treatment options that are accessible and acceptable should be established.

3.7. Facilities for screening/diagnosis and treatment should be available as the lack of follow-up negates the benefit of screening.

3.8. Responsibilities in the screening program should be clear (i.e., who does what and when).

3.9. How the findings will become part of a participant’s medical record should be delineated.

3.10. Compliance with an effective care pathway should be ensured otherwise, there is no benefit of screening.

3.11. Screening programs should be an incessant process rather than being conducted once.

3.12. Continuous monitoring and evaluation should be incorporated into the screening program.

3.13. Consumer perspectives should be integrated.

3.14. Screening programs should not be static but amenable to new scientific evidence.

According to the Ontario Task Force on the Use and Provision of Medical Services (1990), other important questions to consider before developing and implementing a screening program include:

1. Are the screening program requirements (i.e., time and cost) appropriate for the community?

2. Are other equally worthy procedures and efforts being given equivalent consideration or are existing resources being redirected unnecessarily?

3. Does the procedure create new medical risks and how are these assessed in relation to the procedure?

4. Does the procedure place additional strain on health care resources in a disproportionate manner to the magnitude of the health problem being studied?

5. What are the limitations of using screening assessments as a widespread diagnostic tool in relation to other diagnostic approaches?

6. Are there specific ethical or moral issues raised by the screening program?
7. How will the objectives of the screening program be communicated to the various target populations at risk?

*Measures Used in the Detection of Postpartum Depression*

Today, both general and postpartum-specific depression instruments have been utilized to measure depressive symptomatology. The validation of screening tools and the diagnosis of postpartum depression can only be accomplished through the application of *diagnostic criteria* such as the popular and progressively evolving Diagnostic and Statistical Manual [i.e., DSM-III (APA, 1980), DSM-III-R (APA, 1987), or DSM-IV (APA, 1994)] criteria for major depression in addition to the Research Diagnostic Criteria (RDC) (Spitzer, Endicott, & Robins, 1975), and the International Classification of Diseases (ICD-10) (Spitzer et al., 1975; WHO, 1992). Measures used to *assess* depressive symptomatology include standardized interviews, clinician-rated scales, and self-report questionnaires. To provide a clear understanding of the different measures and to promote methodological comparisons between studies, the most common interviews and questionnaires used to assess depressive symptomatology in postpartum depression research are briefly presented.

**Standardized Interviews**

A number of standardized interviews are available to establish a diagnosis of postpartum depression. These instruments are typically used for research purposes and are based on stringent criteria to ensure a systematic and reliable diagnosis. Their use is restricted to trained clinicians or researchers who have a thorough knowledge of DSM, RDC, or ICD systems of diagnosis and clinical judgement is essential in determining whether the responses provided by participants meet the diagnostic criteria. These instruments are time-consuming, expensive, and not recommended for general clinical practice.

*Schedule of Affective Disorders and Schizophrenia (SADS).* The SADS consists of open-ended questions concerning each symptom with probes for follow-up questions (Spitzer, Endicott, & Robins, 1978). There are 11 depressive symptoms (seven somatic and four cognitive affective) in the eight categories of appetite disturbance, sleep disturbance, fatigue, loss of interest, guilt, impaired concentration, suicidal ideation, and motor disturbance. The presence and severity of each symptom is rated from 1 to 6 by the interviewer and a symptom must receive a rating of at least 3 (mild) or higher (severe and experienced often) and have been present for a minimum of 2 weeks to be considered clinically significant. The SADS is designed to obtain data to formulate a diagnosis based on RCD, which has operationally defined inclusion and exclusion criteria for each diagnostic category. Administration takes approximately 90 minutes to complete. The SADS has been used in several postpartum depression studies (Areias, Kumar, Barros, & Figueiredo, 1996a; Carothers & Murray, 1990; Whiffen & Gotlib, 1993).
Structured Clinical Interview for DSM-IV-R (SCID). The SCID is a clinical interview that incorporates DSM-IV diagnoses and has different versions for use with psychiatric inpatient, outpatient, and non-clinical populations (Spitzer, Williams, Gibbon, & First, 1992). While it has software suitable for administration and scoring, clinical judgement is an essential component of the interview, which should be conducted by trained health professionals. It is divided into six self-contained modules and takes approximately 45 to 60 minutes to complete. The SCID has been used in a number of recent postpartum depression studies (Lee et al., 1997; Lee, Yip, Chiu, & Chung, 2000; D. Lee et al., 1998; Zelkowitz & Milet, 1995).

Standard Psychiatric Interview (SPI). The SPI (also referred to as the Clinical Interview Schedule; CIS) is a semi-structured interview intended for use in community surveys (Goldberg, 1972). The SPI is shorter than other standardized interviews and consists of questions designed to elicit the presence or absence of 10 defined psychiatric symptoms. The interviewer rates the presence of another 12 manifest abnormalities of mental state. Each symptom receives a score on a 5-point scale of severity and the total score is the sum of 10 symptom ratings added to twice the score of the manifest abnormalities. The interview has often been modified by adding items concerning appetite changes and weight loss to allow RDC to be applied. The SPI has been used postnatally (Boath, Cox, Lewis, Jones, & Pryce, 1999).

Present State Examination (PSE). The PSE is a semi-structured clinical interview that determines whether or not defined psychiatric symptoms have been present in the previous 4 weeks (Wing & Stuart, 1978). The interview results are used to classify cases according to the PSE-Index of Definition-Category (PSE-ID-Category). The index specifies the degree of certainty with which a respondent may be considered a case, by using eight levels each of which implies greater confidence in case classification; level 5 is considered the threshold that divides cases from non-cases. The criteria used to determine the presence of symptoms are more stringent than are those in the SPI; hence, the SPI could include lower levels of psychiatric morbidity that would not reach the recommended threshold for the PSE. The PSE has been used in a number of postpartum depression studies (Carpiniello, Pariante, Serri, Costa, & Carta, 1997; Ghubash & Abou-Saleh, 1997).

Clinician-Rated Scales

Various clinician-rated scales are available to assess for depressive symptomatology and monitor treatment response. These measures are used to quantify and standardize clinical judgement and provide ratings of duration and severity; they are not employed for population-based screening. The two measures reported most frequently in postpartum depression literature are the Hamilton Rating Scale for Depression and the Montgomery-Asberg Depression Rating Scale.

Hamilton Rating Scale for Depression (HRSD). The HRSD (also referred to as the Hamilton Depression Rating scale - HDRS) was originally developed to assess the severity of depression among diagnosed patients and was intended as a means of qualifying expert clinical judgement (Hamilton, 1960). The original
HRSD consists of 17 depressive symptoms, eight of which relate to somatic complaints, and other versions are available ranging up to 31 items. Responses are rated on either a 3 or 5-point scale with a total score ranging from 0 to 50; a cut-off score of 15 and above is suggestive of major depression. This scale had been used frequently in the postpartum depression literature (Cohen et al., 2001; O’Hara, Stuart, Gorman, & Wenzel, 2000; Thompson, Harris, Lazarus, & Richards, 1998).

Montgomery-Asberg Depression Rating Scale (MADRS). The MADRS was developed as an observer rating scale and consists of 10 items (Montgomery & Asberg, 1979). The items are primarily concerned with psychological symptoms of depression and include global ratings of disturbance and social functioning. Each item is rated in severity from 0 to 6 with a total score ranging from 0 to 60; scores between 7 and 18 indicate mild depression, although some studies have used a cut-off level of 11. While the MADRS has been used by several postpartum depression researchers, it has been associated with a high false positive rate and scores should be confirmed with more reliable methods (Ahokas, Kaukoranta, Wahlbeck, & Aito, 2001; Harris, Johns et al., 1989; Lawrie, Hofmeyr, De Jager et al., 1998; Wickberg & Hwang, 1996b).

Self-Report Questionnaires

Diverse self-report scales are available to assess depressive symptomatology and measure treatment response. These measures generally have respondents rate depressive symptoms in terms of frequency or severity; however, they cannot be used to obtain a diagnosis and high scores should be followed-up with a more in-depth assessment. Some self-report questionnaires are subject to copyright and are not available for general use (e.g., Beck Depression Inventory).

Edinburgh Postnatal Depression Scale (EPDS). By far the most widely used instrument in postpartum depression studies and for population-based screening is the EPDS, a 10-item self-report scale specifically designed to screen for postpartum depression in community samples (Cox, Holden, & Sagovsky, 1987a). Each item is scored on a 4-point scale (from 0 - 3), with a total score ranging from 0 to 30. The items, written in the past tense, include questions related to maternal feelings during the past 7 days and refer to depressed mood, anhedonia, guilt, anxiety, and suicidal ideation. One advantage of this scale is that it does not include common somatic symptoms such as insomnia and appetite changes, which may occur naturally in postpartum women, but rather only one item addresses a somatic symptom and this relates to mood: I have been so unhappy that I have had difficulty in sleeping. The EPDS is typically administered as a pencil and paper test, although computerized versions are now available; both versions are highly correlated and acceptable to women (Glaze & Cox, 1991).

The original EPDS study was completed with a sample of 84 Edinburgh women previously identified by health professionals as potentially depressed at 6 weeks postpartum (Cox et al., 1987a). EPDS scores were compared with the Research Diagnostic Criteria (RDC) obtained from the Standard Psychiatric Interview (SPI). A threshold of 13 identified all 21 women with a RDC diagnosis of major depression and the
sensitivity, specificity, and positive predictive value were 86%, 78%, and 73% respectively. As such, a cut-off score of 12/13 has been recommended for major postpartum depression symptomatology (Cox et al., 1987a; Murray & Carothers, 1990). However, the EPDS does not provide a measure of severity as women who score over 18 can meet DSM criteria for minor depression while others scoring between 14 to 16 can be classified as experiencing major depression (Holden, 1994). Accordingly, the EPDS is not a substitute for a full clinical evaluation but rather a high score is indicative that further assessment is warranted. It is important to note that the selection of a cut-off score depends upon the assessment purpose. While a 12/13 cut-off is suggestive of major depressive symptomatology, a lower threshold of 9/10 has been recommended for community screening to ensure all potential cases of postpartum depression are identified (Cox, Murray, & Chapman, 1993; Murray & Carothers, 1990; Zelkowitz & Milet, 1995).

As shown in Table 1 where studies are chronologically ordered, the English version of the EPDS has been validated in comparison to several standard psychiatric measures (e.g., SADS, SCID, PSE, and SPI) and is highly correlated with other measures of depression including the BDI (Harris, Huckle, Thomas, Johns, & Fung, 1989), SRDS (Condon & Corkindale, 1997), GHQ (Boyce, Stubbs, & Todd, 1993), HRSD (Harris et al., 1992), and MADRS (Harris, Johns et al., 1989). Furthermore, the instrument has been used in various countries resulting in diverse translations and corresponding validation investigations (Table 1). Not surprising, methodological variations, such as population selection criteria, diagnostic criteria, cut-off values, and study timeframe, have resulted in sensitivity and specificity differences. For example, disparities in diagnosis may be problematic as some measures (e.g., PSE) rate depressive symptomatology in the previous 4 weeks while others (e.g., SPI) rate symptoms in the previous 2 weeks; this latter timescale is closer to the EPDS instructions, improving comparison between the two measures. Murray and Carothers (1990) have suggested that sensitivity and specificity may vary according to participants’ ability to identify their psychological status as morbid. They also propose that the EPDS, completed after a semi-structured interview, may not provide the same results as those completed before, as the interview may have sensitised the participant to depressive symptoms that might not have otherwise been acknowledged. Another explanation for the differing sensitivity and specificity is the impact of the reference diagnosis criteria used. For instance, a major depressive diagnosis requires more symptoms to be established in RDC than in DSM. Finally, differences in the positive predictive value are dependent on the prevalence of the condition being examined. Thus, studies with mothers who present clinical symptoms of distress will have a higher prevalence rate and positive predictive value than population-based studies.

These validation studies have also highlighted that scores from translated versions should be interpreted cautiously as different cut-off points have been suggested. For example, Guedeny and Fermanian (1995) concluded in their study that a threshold of 11/12 was appropriate in a French population, giving a sensitivity of 80% and a specificity of 92%. Wickberg and Hwang (1996), validating the EPDS in a Swedish
community sample at 3 weeks postpartum, also suggested an 11/12 cut-off; however, the researchers did not assess the psychometrics using a 9/10 cut-off to provide a true comparison. Ghubask and Abou-Saleh (1999) adopted a threshold score of 11/12 to identify cases of depression among Arabic women when the EPDS was administered at 7 days postpartum and the Present State Examination (PSE) at 8 weeks. Lee et al. (1998) recommended that a cut-off of 9/10 was most appropriate at 6 weeks postpartum in a Chinese population while Okano et al. (1996) reported that a cut-off of 8/9 was suitable for a Japanese population. In an Australian study of Vietnamese and Arabic mothers (Matthey et al., 1997), fewer Vietnamese mothers met the criteria for depression. However, detailed comparisons between EPDS and Diagnostic Interview Schedule (DIS) (Robins, 1989) questions suggested that these lower rates were possibly due to a social desirability bias in terms of verbally reporting negative emotions and a cut-off of 9/10 was suggested for Vietnamese women; similar response patterns were found by Lee et al (1998) in their Hong Kong study. Lee and his colleagues speculated that the traditional supportive rituals of “doing the month” may have postponed the onset of significant depressive symptomatology at 6 weeks postpartum. It is also possible that these Chinese women, like their Vietnamese counterparts, were reluctant to concede unhappiness or distress in the early postpartum period to an interviewer; however, the women seemed less constrained in responding to a self-report questionnaire. In contrast, Yoshida and colleagues (1998) found similar depression rates in Japanese women residing in England and Japan using a clinical interview. However, depression was not detected when the translated EPDS was used as a screening instrument. In particular, a 12/13 cut-off resulted in a sensitivity of zero, rendering the researchers to conclude that Japanese women may be reluctant to disclose depressive symptoms via a self-report measure. They also commented that the difference might be due to the exclusion of somatic symptoms in the EPDS since Japanese women tend to refer to physical problems and concerns about their infant rather than expressing feelings of low mood directly. The preceding results suggest that while an optimal cut-off appears to vary slightly for different cultures, an EPDS score above 9 seems to be the most advantageous threshold if a two-stage screening process (e.g., universal screening where high scoring mothers are contacted further for a more detailed assessment) is implemented to reduce false positive scores.

In addition to widespread usage and sound psychometric properties, the EPDS: (1) is easy to administer, including via telephone (Zelkowitz & Milet, 1995), (2) has uncomplicated interpretation, and (3) can be readily incorporated into routine practice. Furthermore, high maternal acceptance has been reported by numerous researchers (Cox et al., 1987a; Fergusson et al., 2002; Murray & Carothers, 1990; Schaper et al., 1994; Webster et al., 1997; Zelkowitz & Milet, 1995).

**Beck Depression Inventory (BDI).** As one of the most commonly used general self-report questionnaires with considerable psychometric data, including a 25-year review (Beck, Steer, & Garbin, 1988), the copyrighted BDI is a 21-item scale that assesses affect, cognitive symptoms, behaviours, somatic complaints,
and interpersonal domains to measure the presence and intensity of depressive symptoms (Beck, Rush, & Shaw, 1979). Items inquire about mood over the past 7 days and are rated on a 4-point scale ranging from 0 to 3, with higher scores indicating lower mood. While a cut-off score of 12/13 for screening and 20/21 for clinical research has been recommended and many studies have used a cut-off score of 15/16, other researchers have preferred a range of scores with 0 to 9 indicating no symptomatology, 10 to 20 signifying mild depression, 21 to 30 representing moderate depression, and over 30 suggesting severe depression (Kendall, Hollon, & Beck, 1987). Recently, the instrument has been revised to formulate the symptom content to correspond more closely to the diagnostic criteria of DSM-IV. While the BDI-II is still composed of 21 symptoms, the indicators of weight loss, body image change, work difficulty, and somatic preoccupation were eliminated and replaced with the four new symptoms of agitation, worthlessness, concentration difficulty, and loss of energy (Beck, Steer, Ball, & Ranieri, 1996). The scoring is the same as the original BDI but the time period for the ratings has changed from 1 to 2 weeks. The performance of the BDI-II with postpartum women was recently assessed producing acceptable results (Beck & Gable, 2001a).
### Table 2-1. Validation and/or Translation of the Edinburgh Postnatal Depression Scale

<table>
<thead>
<tr>
<th>Study</th>
<th>Year</th>
<th>Country</th>
<th>Language</th>
<th>N</th>
<th>Time</th>
<th>Diagnostic Criteria</th>
<th>Diagnostic Instrument*</th>
<th>EPDS Cut-off</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Cox, Holden, &amp; Sagovsky, 1987b)</td>
<td>1987</td>
<td>UK</td>
<td>English</td>
<td>84</td>
<td>6 wks</td>
<td>RDC</td>
<td>SPI</td>
<td>12/13</td>
<td>86</td>
<td>78</td>
<td>73</td>
</tr>
<tr>
<td>(Harris, Huckle et al., 1989)</td>
<td>1989</td>
<td>UK</td>
<td>English</td>
<td>147</td>
<td>6 wks</td>
<td>DSM-III</td>
<td>Clinical interview, MADRS, BDI</td>
<td>12/13</td>
<td>95</td>
<td>93</td>
<td>75</td>
</tr>
<tr>
<td>(Murray &amp; Cox, 1990)</td>
<td>1990</td>
<td>UK</td>
<td>English</td>
<td>100</td>
<td></td>
<td>RDC</td>
<td>SPI</td>
<td>12/13</td>
<td>64</td>
<td>90</td>
<td>50</td>
</tr>
<tr>
<td>(Carothers &amp; Murray, 1990)</td>
<td>1990</td>
<td>UK</td>
<td>English</td>
<td>646</td>
<td>6 wks</td>
<td>RDC</td>
<td>SPI</td>
<td>9/10</td>
<td>82</td>
<td>89</td>
<td>39</td>
</tr>
<tr>
<td>(Murray &amp; Carothers, 1990)</td>
<td>1990</td>
<td>UK</td>
<td>English</td>
<td>142</td>
<td>6 wks</td>
<td>RDC</td>
<td>SPI</td>
<td>9/10</td>
<td>89</td>
<td>82</td>
<td>39</td>
</tr>
<tr>
<td>(Pop, Komproe, &amp; van Son, 1992)</td>
<td>1992</td>
<td>Netherlands</td>
<td>Dutch</td>
<td>303</td>
<td>4 wks</td>
<td>—</td>
<td>SCL-90, BDI</td>
<td>Correlations with other depression scales</td>
<td>9/10</td>
<td>100</td>
<td>89.4</td>
</tr>
<tr>
<td>(Boyce et al., 1993)</td>
<td>1993</td>
<td>Australia</td>
<td>English</td>
<td>103</td>
<td>12 wks</td>
<td>DSM-III-R</td>
<td>DIS, GHQ, Pitt</td>
<td>9/10</td>
<td>100</td>
<td>95.7</td>
<td>69.2</td>
</tr>
<tr>
<td>(Jadresic, Araya, &amp; Jara, 1995)</td>
<td>1995</td>
<td>Chile</td>
<td>Spanish</td>
<td>108</td>
<td>8-12 wks</td>
<td>RDC</td>
<td>PAS</td>
<td>9/10</td>
<td>100</td>
<td>80</td>
<td>37</td>
</tr>
<tr>
<td>(Zelkowitz &amp; Milet, 1995)</td>
<td>1995</td>
<td>Canada</td>
<td>English</td>
<td>89</td>
<td>6-8 wks</td>
<td>DSM-III-R</td>
<td>SCI</td>
<td>9/10</td>
<td>91</td>
<td>76</td>
<td>78</td>
</tr>
<tr>
<td>(Wickberg &amp; Hwang, 1996b)</td>
<td>1996</td>
<td>Sweden</td>
<td>Swedish</td>
<td>128</td>
<td>8-12 wks</td>
<td>DSM-III-R</td>
<td>MADRS</td>
<td>11/12</td>
<td>96</td>
<td>49</td>
<td>—</td>
</tr>
<tr>
<td>(Areias et al., 1996a)</td>
<td>1996</td>
<td>Portugal</td>
<td>Portuguese</td>
<td>54</td>
<td>24 wks</td>
<td>RDC</td>
<td>SADS</td>
<td>9/10</td>
<td>65</td>
<td>96</td>
<td>91</td>
</tr>
<tr>
<td>(Okano et al., 1998)</td>
<td>1996</td>
<td>Japan</td>
<td>Japanese</td>
<td>—</td>
<td></td>
<td>—</td>
<td>—</td>
<td>8/9</td>
<td>75</td>
<td>93</td>
<td>50</td>
</tr>
<tr>
<td>(Ghubash, Abou-Saleh, &amp; Daradkeh, 1997)</td>
<td>1997</td>
<td>UAE</td>
<td>Arabic</td>
<td>93</td>
<td>1wk EPDS 8wk PSE</td>
<td>PSE-ID-Catego</td>
<td>PSE</td>
<td>10/11</td>
<td>91</td>
<td>84</td>
<td>44</td>
</tr>
<tr>
<td>(Guedeney, Fermanian, Guelfi, &amp; Kumar, 2000)</td>
<td>2000</td>
<td>France</td>
<td>French</td>
<td>87</td>
<td>16 wks</td>
<td>RDC</td>
<td>PSE</td>
<td>9/10</td>
<td>84</td>
<td>78</td>
<td>30</td>
</tr>
<tr>
<td>(Carpiniello et al., 1997)</td>
<td>1997</td>
<td>Italy</td>
<td>Italian</td>
<td>61</td>
<td>4-6 wks</td>
<td>PSE-ID-Catego</td>
<td>PSE</td>
<td>9/10</td>
<td>100</td>
<td>83</td>
<td>50</td>
</tr>
<tr>
<td>(D. T. Lee et al., 1998)</td>
<td>1998</td>
<td>Hong Kong</td>
<td>Chinese</td>
<td>142</td>
<td>6 wks</td>
<td>DSM-III-R</td>
<td>SCI, GHQ, BDI</td>
<td>9/10</td>
<td>82</td>
<td>86</td>
<td>44</td>
</tr>
</tbody>
</table>

*C. Cindy-Lee Dennis, PhD*
<table>
<thead>
<tr>
<th>Study</th>
<th>Year</th>
<th>Country</th>
<th>Language</th>
<th>N</th>
<th>Time</th>
<th>Diagnostic Criteria</th>
<th>Diagnostic Instrument*</th>
<th>EPDS Cut-off</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Bergant, Nguyen, Heim, Ulmer, &amp; Dapunt, 1998)</td>
<td>1998</td>
<td>Austria</td>
<td>German</td>
<td>110</td>
<td>4 days</td>
<td>ICD-10</td>
<td>Clinical interview</td>
<td>9/10</td>
<td>96</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>(Lawrie, Hofmeyr, de Jager, &amp; Berk, 1998)</td>
<td>1998</td>
<td>South Africa</td>
<td>English</td>
<td>102</td>
<td>6 wks</td>
<td>DSM-IV</td>
<td>Clinical interview, MADRS</td>
<td>9/10</td>
<td>84</td>
<td>57</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>12/13</td>
<td>76</td>
<td>82</td>
<td>58</td>
</tr>
<tr>
<td>(Clifford, Day, Cox, &amp; Werrett, 1999)</td>
<td>1999</td>
<td>UK</td>
<td>Punjabi</td>
<td>98</td>
<td>6-8 wks</td>
<td>—</td>
<td>—</td>
<td>Conceptual and cross-cultural equivalence</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Benvenuti, Ferrara, Niccolai, Valoriani, &amp; Cox, 1999)</td>
<td>1999</td>
<td>Italy</td>
<td>Italian</td>
<td>113</td>
<td>8-12 wks</td>
<td>DSM-III-R</td>
<td>MINI</td>
<td>9/10</td>
<td>83</td>
<td>90</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>12/13</td>
<td>56</td>
<td>99</td>
<td>91</td>
</tr>
<tr>
<td>(Barnett, Matthey, &amp; Gyaneshwar, 1999)</td>
<td>1999</td>
<td>Australia</td>
<td>English</td>
<td>105</td>
<td>6 wks</td>
<td>DSM-III-R</td>
<td>DIS, GHQ-30, Faces Scale</td>
<td>9/10</td>
<td>100</td>
<td>69</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Arabic</td>
<td>98</td>
<td>6 wks</td>
<td>DSM-III-R</td>
<td>DIS, GHQ-30, Faces Scale</td>
<td>9/10</td>
<td>78</td>
<td>80</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Vietnamese</td>
<td>113</td>
<td></td>
<td>DSM-III-R</td>
<td>DIS, GHQ-30, Faces Scale</td>
<td>12/13</td>
<td>56</td>
<td>91</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>9/10</td>
<td>86</td>
<td>84</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>12/13</td>
<td>57</td>
<td>94</td>
<td>40</td>
</tr>
<tr>
<td>(Thome, 2000)</td>
<td>2000</td>
<td>Iceland</td>
<td>Icelandic</td>
<td>201</td>
<td>8-12 wks</td>
<td>—</td>
<td>—</td>
<td>Cronbach’s alpha 0.80</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Eberhard-Gran, Eskild, Tambs, Schei, &amp; Opjordsmoen, 2001)</td>
<td>2001</td>
<td>Norway</td>
<td>Norwegian</td>
<td>56</td>
<td>8-12 wks</td>
<td>DSM-IV</td>
<td>PCEMD, MADRS</td>
<td>9/10</td>
<td>100</td>
<td>87</td>
<td></td>
</tr>
<tr>
<td>(Regmi, Sliq, Carter, Grut, &amp; Seear, 2002)</td>
<td>2002</td>
<td>Nepal</td>
<td>English</td>
<td>100</td>
<td>8-12 wks</td>
<td>DSM-IV</td>
<td>Structured interview</td>
<td>12/13</td>
<td>100</td>
<td>93</td>
<td>42</td>
</tr>
</tbody>
</table>

* SCL-90 – Symptom Checklist-90; PAS-Psychiatric Assessment Schedule; MINI – Mini International Neuropsychiatric Interview; DIS – Diagnostic Interview Schedule; PCEMD – Primary Care Evaluation of Mental Disorders; SCI – Structured Clinical Interview for DSM-III-R
While the BDI’s psychometric properties have demonstrated robustness as an instrument, its use as a postpartum depression measure is equivocal (Hopkins, Campbell, & Marcus, 1989; Horowitz, Damato, Solon, Von Metzsch, & Gill, 1995) as several studies have found the instrument to be unsatisfactory as a screening measure (Gotlib, Whiffen, Mount, Milne, & Cordy, 1989; Whiffen, 1988). In particular, the large number of somatic items, which are normal postpartum symptoms, have lead to inflated scores among new mothers (Harris, Huckle et al., 1989; Hopkins et al., 1989; O'Hara, Neunaber, & Zekoski, 1984). Furthermore, in a Dutch population, 12% of mothers expressed difficulty in their ability to complete the BDI (Pop et al., 1992). Despite these cautions, researchers have suggested that the BDI is valuable in studies involving longitudinal designs and measurement of symptom severity (Affonso et al., 2000).

Center for Epidemiological Studies Depression Scale (CES-D). The CES-D consists of 20 items chosen from previously validated depression scales with an emphasis on the affective component of depressed mood (Radloff, 1977). Items inquire about mood in the past 7 days and are rated on a 4-point scale with scores ranging from 0 and 60, with higher scores indicating lower mood. Sixteen items represent negative symptoms such as depressed mood, feelings of guilt, and worthlessness and helplessness, whereas four positively worded items are included to break tendencies and assess positive affect and sense of well-being (Liang, Van Tran, Krause, & Markides, 1989). These four items are reverse coded to indicate lack of well-being. A score of 16 has been used as a standard threshold indicating possible clinical depression (Radloff, 1977; Weissman, Sholomskas, Pottenger, Prusoff, & Locke, 1977). The CES-D has been used in several postpartum depression studies (Campbell & Cohn, 1991; Fleming, Klein, & Corter, 1992; Logsdon, McBride, & Birkimer, 1994) and in a sample of 1,007 primiparous women, the CES-D had a sensitivity of 60% and a specificity of 92% (Campbell & Cohn, 1991).

Depression Adjective Checklist (DACL). The DACL has seven equivalent checklists to minimise test-retest effects and all contain either 32 or 34 adjective choices (Lubin, 1981; Lubin, Nathan, & Nathan, 1981). Negative (e.g., weary or low spirited) and positive (e.g., joyous or enthusiastic) adjectives are listed and respondents check all the words that describe how they feel ‘now-today.’ Most items are expressed in terms of affect rendering the DACL to be regarded more a measure of depressed mood than indicative of a depressive syndrome. Scores range from 0 to 32 or 34, depending on the form used, with higher scores indicating increased depressed affect. Limitations of this scale include its failure to evaluate duration or severity of dysphoria and that it has not been specifically validated with pregnant or postpartum women. As such, it is seldom used in postpartum depression studies (Da Costa, Larouche, Dritsa, & Brender, 2000; Gennaro, 1988; Horowitz, Damato, Solon, & von Metzsch, 1996; Pop et al., 1992).

General Health Questionnaire (GHQ). The GHQ covers a broad range of symptoms related to psychiatric disorders among a general population and is divided into four subscales: Somatic, Anxiety and Insomnia, Social Dysfunction, and Severe Depression (Goldberg & Hillier, 1979). Each item is rated on a 4-point
Likert-type format ranging from 0 (absent) to 3 (intense). There are several versions of the GHQ each containing different numbers of items ranging from 60 (GHQ-60) to 12 (GHQ-12) items. While the GHQ has been used in several postpartum depression studies (Boyce et al., 1993; Guedeney & Fermanian, 1998; D. Lee et al., 1998; Matthey et al., 1997), new mothers frequently have inflated scores. For example, in an Australian study of 103 postpartum women, 25% scored over 4, the standard cut-off point (Boyce et al., 1993). Using a slightly modified version of the GHQ-30 (removing two questions pertaining to disturbed sleep and getting out of the house) and raising the cut-off to over 6 improved the GHQ as a measure of postpartum depression (Brugha et al., 1998).

**Hospital Anxiety and Depression Scale (HADS).** The HADS is 14-item scale that contains seven items pertaining to depressed mood (HADS-D) (Zigmond & Snaith, 1983). Scores range between 0 and 21 and although it includes one item relating to feeling ‘slowed down,’ it otherwise excludes neurovegetative changes. While data about sensitivity and specificity is not provided, a cut-off score of 11 has been equated with depressive symptomatology and the scale has been used in several postpartum depression studies (Condon & Corkindale, 1997; Thompson et al., 1998). However, in a study of 755 women, the HADS-D had a sensitivity of 65% and specificity of 90% resulting in the researchers suggesting that a lower cut-off of 8/9 should be employed in postpartum samples (Harris et al., 1992); similar results were found by others (Thompson et al., 1998).

**Profile of Mood States (POMS).** The POMS is a 65-item, adjective-rating scale designed to measure subjective mood states where respondents are presented with a list of feelings and requested to reply to the question "How have you been feeling during the past week including today?" (McNair, Loot, & Droppleman, 1981). Each question is rated on a 5-point Likert-type scale ranging from 0 (not at all) to 4 (extremely) with a total score obtained by summing the 58 items on the following factors: Tension-Anxiety, Depression-Dejection, Anger-Hostility, Vigour (reverse scored), Fatigue, and Confusion; the remaining seven items are included as buffers. The depression subscale (POMS-D) contains 15 items, including somatic ones. Recommended cut-off scores have not been established although higher scores indicate increased mood disturbance. While it has not been widely used in the general depression literature, it has been incorporated in a number of studies with postpartum women (Condon & Corkindale, 1997; Fisher, Feekery, & Rowe-Murray, 2002; Hayes, Muller, & Bradley, 2001; Meager & Milgrom, 1996).

**Pitt Depression Scale.** The Pitt Depression Scale is a 24-item questionnaire designed as a screening instrument to measure maternal anxiety and depression before and after childbirth (Pitt, 1968). The items are listed as questions and the respondent indicates whether each symptom was present ‘today, or over the past few days’ and are given a choice of responding yes, no, or don’t know; total scores range from 0 to 48. This scale has been used infrequently with postpartum women (Boyce et al., 1993; Wolman, Chalmers, Hofmeyr, & Nikodem, 1993).

*Cindy-Lee Dennis, PhD*
Postpartum Depression Screening Scale (PDSS). The PDSS is a 35-item Likert-type response scale consisting of seven dimensions, each of which contains five items; the dimensions include sleeping/eating disturbances, anxiety/insecurity, emotional lability, cognitive impairment, loss of self, guilt/shame, and contemplating harming oneself (Beck & Gable, 2000, 2001b). The conceptual basis of the PDSS is based on a series of qualitative postpartum depression studies (Beck, 1992, 1993, 1996). Each item describes how a woman may be feeling after the birth of her baby and respondents are asked to indicate their degree of disagreement or agreement on a 5-point scale regarding how they have felt over the past 2 weeks. Initial psychometric testing involved 525 mothers at approximately 6 weeks postpartum. In a further methodological study, 150 mothers within 12 weeks postpartum completed in random order three questionnaires: PDSS, EPDS, and the Beck Depression Inventory II (BDI-II). The PDSS was strongly correlated with both the BDI-II ($r = 0.81$) and EPDS ($r = 0.79$). Using Receiver Operating Characteristic (ROC) curves, a PDSS cut-off score of 80 (sensitivity 94% and specificity 98%) was recommended for major postpartum depression and a cut-off score of 60 (sensitivity 91% and specificity 72%) for major or minor postpartum depression.

Zung Self-Rating Depression Scale (ZSDS). The ZSDS is a widely used and extensively validated self-report instrument containing 20 items: three affect items, six cognitive items, four overt-motor behaviour items, six somatic items, and one social-interpersonal item (Zung, Richards, & Short, 1965). Respondents rate each item according to how they felt during the preceding week with item responses ranked from 1 to 4, with higher numbers corresponding to more frequent symptoms (although several items are scored in reverse). The sum of the 20 items produces a raw score that is converted into a percentage of the depression measurable by the scale (termed the “SDS Index”). Index scores are then categorized into 4 levels to offer a global clinical impression, as recommended by the instrument developers: I, within normal range, no significant psychopathology (SDS Index: <50); II, presence of minimal to mild depression (SDS Index: 50 - 59); III, presence of moderate to marked depression (SDS Index: 60 - 69); and IV, presence of severe to extreme depression (SDS $\geq 70$). While the ZSDS has been used in diverse postpartum depression studies in a variety of countries (Augusto, Kumar, Calheiros, Matos, & Figueiredo, 1996; Condon & Corkindale, 1997; Kitamura, Shima, Sugawara, & Toda, 1994; Viinamaki, Niskanen, Pesonen, & Saarikoski, 1997), limitations such as length and copyright limit feasibility.

Comparisons between Screening Instruments

It has been suggested that the measurement of ‘depression’ is as confused as the basic construct itself. Other researchers have summarized the inherent difficulties of assessing the presence and/or severity of psychiatric syndromes from rating scales based on symptoms (Snaith, 1993). Aside from the problems of agreeing upon an appropriate cut-off score for diagnostic caseness, there is the more fundamental difficulty of a lack of agreement regarding the definition of depression. If a construct such as ‘depression’ is defined
very narrowly, items contained in a scale to measure it will tend to be homogeneous and may paraphrase each other. Such an instrument will have high internal consistency but may lack validity. In contrast, if depression is defined too broadly, then quite different symptom profiles could achieve similar scores. As a result, diagnostic caseness may be achieved by disparate groups of individuals. The problem is how much heterogeneity is acceptable if the construct depression is to remain useful and meaningful (Condon & Corkindale, 1997). The weight of available evidence supports the notion that depression is heterogeneous; therefore, the validity of a particular questionnaire score may be debatable if the instrument is used in populations different from the one with which it was developed. It has also been claimed that the fundamental lack of agreement regarding the definition of depression and appropriate cut-off scores for diagnosis result in low levels of concurrent validity between different measurements (Condon & Corkindale, 1997).

Diversity and inconsistency in assessment procedures for postpartum depression have hampered the meaningful comparison of studies and compromised the development of a cumulative body of knowledge. The lack of consensus regarding the definition of postpartum depression has resulted in different types of assessment instruments, variable cut-off scores, and diverse times in which assessments are conducted. In order to attain standardization, systematic scrutiny regarding the advantages and disadvantages of assessment methods used is required. For example, self-report measures have the advantage of being relatively inexpensive and easy to use. Furthermore, administration of these measures requires little time or previous training, which permits wider use than clinician-rated scales or standardized interviews. However, self-report measures do not incorporate the benefit of clinical judgement in the weighing of symptoms or enquiry about the context of symptoms such as sleep disturbances.

Despite these conceptual and methodological issues, several researchers have conducted comparisons between diverse self-report measures to determine which instrument is the most effective in identifying postpartum mothers with depressive symptomatology. In a UK study, 147 mothers were screened for major depression at 6 to 8 weeks postpartum. Using predetermined cut-off points, the EPDS and BDI were compared in their abilities to identify the 15% of mothers who were diagnosed with major depression according to DSM-III criteria (Harris, Huckle et al., 1989). The sensitivity of the EPDS was 95% and its specificity 93%. The performance of the Beck Depression Inventory (BDI) was markedly inferior, with a sensitivity of 68% and specificity of 88%. Similarly, the results of a study looking into the association between thyroid status and postpartum depression were reanalysed to explore the psychometric properties of the rating scales employed (Thompson et al., 1998). The performance of the EPDS was found to be superior to that of the Hospital Anxiety and Depression Scale (HADS) in identifying RDC-defined depression and on par with the observer-rated Hamilton Rating Scale for Depression (HRSD). In an Australian study, 200 mothers completed questionnaires at 4, 18, and 32 weeks postpartum to ascertain the degree of agreement...
between four self-report depression scales, with particular emphasis on whether each scale would identify the same subgroup of women as being 'most depressed' (Condon & Corkindale, 1997). The four instruments included were the EPDS, the depression subscale of the Hospital Anxiety Depression Scale, the Zung Self-Rating Depression Scale, and the depression subscale of the Profile of Mood States. Agreement between pairs of instruments, in terms of identifying the most depressed subgroup of women in the cohort, averaged approximately 40%; agreement between the three instruments was only about 25%. This poor level of agreement most likely reflects the different emphasis in item content of the questionnaires, which in turn clearly signals the distinct notions of 'depression' held by the instrument developers.

To compare the performance of the newly created Postpartum Depression Screening Scale (PDSS) with the EPDS and Beck Depression Inventory-II (BDI-II), 150 US women completed these instruments in random order, followed immediately by a DSM-IV diagnostic interview (Beck & Gable, 2001a). Of the 150 participants, 18 (12%) were diagnosed with major postpartum depression, 28 (19%) with minor postpartum depression, and 104 (69%) with no depression. The areas under each of the instrument's Receiver Operating Characteristic (ROC) curves were compared to determine significant discrepancies. A ROC curve is constructed by plotting the sensitivity (i.e., true-positive rate) against the false positive rate (i.e., rate at which an instrument falsely indicates the presence of postpartum depression in non-depressed mothers) over a range of cut-off scores (Fletcher, Fletcher, & Wagner, 1996). The overall accuracy of an instrument can be described as the area under the ROC curve. The larger the area is under the curve (AUC), the better the classification ability of the instrument. Compared to the EPDS, the PDSS had a significantly larger area under the ROC curve when screening for major or minor postpartum depression. When using published recommended cut-off scores for major depression, the PDSS achieved the highest combination of sensitivity (94%) and specificity (98%). When detecting women with major or minor postpartum depression, the PDSS again yielded the highest combination of sensitivity (91%) and specificity (72%) of the three instruments. The PDSS identified 17 (94%) of the women diagnosed with major postpartum depression, the EPDS identified 14 women (78%), and the BDI-II identified 10 women (56%). These results are promising and further research with the PDSS is warranted, however its cost and length are barriers to wider use.

In another psychometric study, Chinese women completed the General Health Questionnaire (GHQ), Beck Depression Inventory (BDI), and EPDS at 6 weeks postpartum and were then assessed using the Structured Clinical Interview for DSM-III-R (SCID) (Lee, Yip, Chiu, Leung, & Chung, 2001). The psychometric performance of the GHQ, BDI, and EPDS in detecting postpartum depression was assessed using the Receiver Operating Characteristic (ROC) curves. Both the Chinese GHQ and BDI had satisfactory sensitivity and positive predictive value in detecting postpartum depression and the ROC curves were comparable to that of the EPDS. While the GHQ and BDI may be useful for detecting postpartum depression
among recently delivered Chinese women, the study was conducted using the translated versions of the rating scales, limiting generalizability to English speaking populations.

Finally, to determine whether applying two complementary rating scales of depression symptomatology as a double test would significantly enhance the positive predictive value of postpartum depression screening, 145 Chinese women completed the EPDS and 12-item General Health Questionnaire (GHQ) at 6 weeks postpartum; clinical interviews were then completed to validate postpartum depression diagnoses (Lee et al., 2000). The positive predictive value of the EPDS and GHQ, when administered independently, was 44% and 52%, respectively, for probable major postpartum depression. When the EPDS-GHQ double test was administered, the positive predictive value was increased significantly to 78%. This preliminary finding suggests that simultaneous administration of the EPDS and GHQ may improve identification of women with postpartum depression, potentially enhancing the overall effectiveness of population-wide screening.

Antenatal Screening

While determining the most appropriate instrument to detect postpartum depression is exigent, immense efforts have also been undertaken to identify pregnant women who are at-risk of developing postpartum depression such that secondary preventive interventions may be implemented (Table 2-2). One of the earliest studies to design a simple and practical instrument to detect high-risk women was conducted in Canada (Braverman & Roux, 1978). Randomly selected women (N = 120) attending a Montreal-based prenatal clinic were requested to complete a 19-item "yes/no" questionnaire. Each mother was classified for presence or absence of “postpartum emotional disorder” (PED), according to clearly defined criteria. The responses of mothers classified as having emotional disorders (13%) were compared to the "normal" group with 6 items showing predictive value: (1) feeling unloved by husband, (2) feeling the pregnancy was undesired, (3) past history of postpartum depression, (4) being single or separated, (5) marital problems, and (6) unplanned pregnancy. Following this pioneering work, 17 studies have been found assessing the performance of antenatal questionnaires in predicting postpartum depression (Table 2-1). Chronologically, the Leverton Questionnaire was derived from five factors associated with postpartum depression including past psychiatric history, maternal anxiety, dissatisfaction with marital relationship, lacking a confidante, and previous history of postpartum depression (Leverton & Elliott, 1988). Also included were items pertaining to socio-demographic variables, feelings about pregnancy, previous obstetric and gynaecological history, current stressors, and depression and somatic items from the Crown Crisp Experimental Index (Crown & Crisp, 1979). The questionnaire was validated with 188 pregnant UK women having their first or second child with 99 (53%) being identified as “more vulnerable” for postpartum depression. Fifty of these women were then compared with “less vulnerable” women (n = 89) at 12 weeks postpartum using a detailed interview, including the Present State Examination (PSE). In the vulnerable group, 20 (40%) women were identified with definite/borderline postpartum depression in comparison to only 14 (16%) women without vulnerable
factors. However, the questionnaire identified over half of the sample as vulnerable suggesting a low specificity and positive predictive value.

In a sample of 192 financially impoverished, inner-city women, clinical depression was assessed twice antenatally and once postnatally (Hobfoll, Ritter, Lavin, Hulsizer, & Cameron, 1995). Using the Schedule for Affective Disorders and Schizophrenia (SADS) clinical interview and after controlling for pregnancy-related somatic symptoms, 27.6% of women were identified as depressed at the first antenatal interview, decreasing to 24.5% at the second assessment. Postpartum depression was found among 23.4% of women, a rate significantly higher than those found in middle-class samples. A particularly salient risk factor for antepartum depression was single status while depression rates did not differ between African American and European American women. It is noteworthy that antepartum depression was a weak risk factor for postpartum depression.

In another UK study, a 10-item screening questionnaire was constructed from previous reports of postpartum depression risk factors (Appleby, Gregoire, Platz, Prince, & Kumar, 1994). The predictive ability of the tool was tested among 165 women attending an antenatal clinic at 36 weeks gestation who were then assessed for postpartum depression using the EPDS at 8 weeks postpartum. One hundred and twenty-six (77%) mothers returned the EPDS, 13% of whom had a score above 11. Neither the Antenatal Screening Questionnaire as a whole, nor groups of items, was able to discriminate well between women who later developed depressive symptomatology and the predictive ability of the questionnaire accounted for only 6% of the variance in EPDS scores. Although the antenatal questionnaire scores weakly correlated with postpartum EPDS scores, this was largely because the questionnaire was able to identify correctly those who would not become depressed.

In a study comparing correlates of paternal and maternal depression, 54 Portuguese primiparous mothers attending obstetric services participated in a longitudinal study of their mental health (Areias, Kumar, Barros, & Figueiredo, 1996b). All mothers were given a semi-structured clinical interview (SADS) at 24 weeks antenatally and 52 weeks postnatally and sub-samples were interviewed at 12 weeks postpartum. At these time periods, mothers also completed a translated version of the EPDS. Aside from a history of depression, the only other significant predictor of postpartum depression was negative life events.

Based on the Leverton Questionnaire (Leverton & Elliott, 1988), the Modified Antenatal Screening Questionnaire (MASQ) was developed to identify women vulnerable to become depressed after childbirth (Stamp, Williams, & Crowther, 1996). Two hundred and forty nine Australian women at 24 weeks gestation or less completed the screening questionnaire of which 144 (58%) screened more vulnerable; at 6 weeks postpartum, participants completed the EPDS. No difference was found at 6 weeks postpartum between the vulnerable group (return rate 64/68) and the less vulnerable group (return rate 44/51) in the frequency of those who screened as high-risk for postpartum depression. For probable major postpartum depression
In a sample of over 6000 UK women recruited in the last trimester of pregnancy, a 40-item self-report questionnaire designed to detect risk factors for postpartum depression was administered and maternal mood was assessed at 6 to 8 weeks postpartum using the EPDS (Cooper, Murray, Hooper, & West, 1996). A total of 5,124 (86%) women completed the EPDS at 5 to 6 weeks of which 1,629 scored above 8 and 1,459 (90%) were contacted again by telephone to assess depressed mood and anhedonia. If both factors were not denied, women were interviewed at home with the SCID to establish DSM-III-R diagnosis of major depression. Through a series of logistic regressions on two-thirds of this sample, the original set of variables was reduced to a predictive index of 17 items with weighted scores calculated for each. This index was then applied to the remaining one-third of the sample as a validating procedure and specificity and sensitivity was calculated. The overall rate of major postpartum depression was 15.3%. To determine the predictive performance of the index, they assumed that at a base postpartum depression rate of 10% to 15%, a score of 35 or more resulted in a 40% risk of developing depression; however, 95% of those who were to become depressed scored below 35 on the index. At a score of 27 or more, the risk of postpartum depression was 35% with more than a third of those who were to become depressed scoring in this range. While the researchers recommended the clinical use of the index and the study was well designed with a large sample size, the poor predictive power does not support the utility assertions. It is noteworthy that the researchers suggested that the predictive performance would be significantly improved if maternity blues and infant factors were included.

In another study to develop an antenatal questionnaire, demographic and clinical data, based on previously identified variables, were obtained from 106 pregnant women in their second-trimester (sample I) (Posner, Unterman, Williams, & Williams, 1997). The Beck Depression Inventory (BDI) was then administered at 1, 6, and 12 weeks postpartum. Statistical analysis, including stepwise linear regression, identified a subset of 24 predictive variables. This Antepartum Questionnaire (APQ) was validated retrospectively in the original sample and prospectively in a second group of 99 women (sample II). In both samples, the APQ had acceptable sensitivity (80-82%) and specificity (78-82%). The incidence of postpartum depressive symptoms rose from 10% to 17% by 6 weeks without an appreciable decline at 12 weeks (15%). The percentage of women with more than mild depressive symptoms increased from 30% at 1 week to 47% at 12 weeks. However, the high number of women identified as high-risk and the low positive predictive value suggests the APQ also has limited clinical utility.
using the Beck Depression Inventory (BDI) and at 6 weeks postpartum using the EPDS (Glasser et al., 1998). While the two-thirds of the cohort scored as ‘depressed’ antenatally, one-third of the mothers identified with postpartum depression were new cases. Immigrant status was the only significant predictor of postpartum depression, with Russian new immigrants having over twice the risk for postpartum depression as Israeli-born mothers.

In a prospective cohort study conducted in Zimbabwe, 500 women in the eighth month of pregnancy identified by traditional birth attendants and primary care clinics completed the Shona Symptom Questionnaire (SSQ), a 14-item indigenous psychiatric questionnaire based on local idioms focusing on cognitive symptoms (Nhiwatiwa, Patel, & Acuda, 1998). A “high-risk” cohort consisted of all women who scored above 7 on the questionnaire (n = 95) and a “low-risk” cohort of 105 women was randomly selected from the remainder of the sample; a modified Clinical Interview Schedule (CIS) was completed by all participants at 6 to 8 weeks postpartum with a score of 14 or more indicating psychiatric caseness. In this study, the overall prevalence of postpartum ‘mental illness’ was 16%. Of the 95 high-risk women, 44 (46%) scored above 13 on the CIS in comparison to only 10 (9%) of the low-risk mothers. The odds ratios for high-risk women to become mentally unwell in the postpartum period was 10.6 (95% CI = 4.8 - 23.9, p < 0.001) after adjusting for age, marital status, and occupation. The most serious study limitation was that a categorical approach was not used in classifying postpartum mental disorders into various types. Even though the researchers suggested that depressive symptoms were the most common clinical presentation among women classified with a “postpartum mental disorder,” it is unknown how many women were actually experiencing postpartum depression.

In a pilot trial, pregnant Australian women between 12 and 24 weeks gestation were screened for postpartum depression risk factors based on a researcher-developed ‘risk factor scale’ (Buist, Westley, & Hill, 1999). Risk factors were selected from a literature review with women scoring above 7 identified as ‘at-risk;’ a score on this scale reflected a mix of three or more risk factors related to personal/family history of depression, premenstrual syndrome, and marital/childhood difficulties. While 23% of all screened mothers (N = 348) were identified as high-risk, no sensitivity or specificity results were reported and no mother recruited to participate in the preceding pilot trial (N = 44) had an EPDS score above 12 at any point in time, providing little support for the screening tool.

In another controlled trial to evaluate the effect of an antenatal intervention to prevent postpartum depression (Brugha et al., 2000), 1300 primiparous UK women were screened using a “Pregnancy and You” questionnaire; the presence of one of the six depression items from the modified General Health Questionnaire resulted in a positive screen. Thirty-one percent of women screened positive and while limited data is available on the screening properties, the low predictive value is evident (Lumley & Austin, 2001).
In a similar trial evaluating a preventive intervention (Elliott et al., 2000), UK women expecting their first or second child completed the Leverton Questionnaire (Leverton & Elliott, 1988) and the depression, anxiety, and somatic subscales of the Crown Crisp Experiential Index (CCEI) (Crown & Crisp, 1979); two questions from a Canadian questionnaire (Braverman & Roux, 1978) were also included. Vulnerability items were constructed to represent four factors: (1) dissatisfaction with partner, (2) previous psychiatric history, (3) lacking a confidante, and (4) high antenatal anxiety. Women were classified as vulnerable if they scored 2 on the Leverton Questionnaire, or scored 1 on more than one question. Women who scored 10 or higher on the CCEI or had a previous history of feeling tense or depressed after childbirth were also considered vulnerable. Obviously, this screening tool is neither simple nor clinically practical and the 38% positive predictive value led the researchers to comment, “It seems unlikely that an antenatal screening questionnaire for postnatal depression could be produced with sufficient predictive power to be clinically useful.”

In a population-based study designed to test the predictive power of demographic, obstetric, and psychosocial risk factors related to postpartum depression, 6790 Danish women who attended an antenatal clinic were recruited with 5252 (78%) completing all questionnaires (Forman, Videbech, Hedegaard, Salvig, & Secher, 2000). The validation population was comprised of a separate sample of 528 women enrolled immediately before and after the study period. While more than a third of all pregnant women were identified as being at high risk, only 5.5% of the women scored above 12 on the EPDS. While the sensitivity was high, one in five women who developed depressive symptomatology were missed antenatally and only 12% of mothers identified as high risk went on to develop postpartum depression. Even though this study was well conducted and the tool was validated with a separate sample, the unexplainably low postpartum depression prevalence rate resulted in a low positive predictive power limiting clinical utility.

In a prospective longitudinal study conducted in the UK, 417 pregnant women completed the EPDS antenatally and at 12 weeks postpartum (Johanson, Chapman, Murray, Johnson, & Cox, 2000). Using an unusually high cut-off score (above 14), 41 (9.8%) women during pregnancy and 31 (7.4%) at 12 weeks postpartum were identified with depressive symptomatology. While there was a significant association between antenatal and postpartum depressive symptomatology, only seven (22.6%) of the 31 women who were depressed postnataally had also been depressed antenatally. The unacceptably low sensitivity and predictive power suggests the EPDS is also a poor antenatal screening tool.

In an Australian study of 2118 pregnant women, 901 women (600 with and 301 without antenatal risk factors for postpartum depression) were recruited and administered an antenatal screening tool with 574 (86.4%) returning a postpartum EPDS at 16 weeks (Webster, Linnane, Dibley, & Pritchard, 2000). While more women (25.9%) with an antenatal risk factor scored above 12 on the EPDS than those without any risk factor (10.9%) (p < 0.001), 40% of women who scored 3 or more on the “Postnatal Depression Risk Index” experienced postpartum depression representing only 27% of all women scoring over 12 on the EPDS at 16
weeks postpartum. Using a cut-off score of 2 on the questionnaire increased the proportion of depressed women correctly identified to 44%; however, the risk of postpartum depression among women who scored in this range fell to 32% (positive predictive value). While the researchers suggest their screening process was effective, their results suggest otherwise.

To examine the prevalence of depressive symptoms and determine whether there is an association between antenatal and postnatal depressive symptomatology, a longitudinal study of 1,558 pregnant Swedish women was conducted (Josefsson, Berg, Nordin, & Sydsjo, 2001). The presence of depressive symptoms was measured using EPDS on four occasions: 35 to 36 weeks gestation, immediately post-delivery, 6 to 8 weeks postpartum, and 24 weeks postpartum; respective prevalence rates were 17, 18, 13, and 13%. While a correlation between antenatal and postnatal depressive symptoms was found ($r = 0.50$, $p<0.0001$), the positive predictive value was only 33%.

In another postpartum depression preventive trial, 135 pregnant women receiving public assistance were screened for at least one risk factor for postpartum depression; 67% of women screened at risk resulting in 37 women being randomly assigned to either a four-session interpersonal psychotherapy group intervention or to a treatment-as-usual condition (Zlotnick, Johnson, Miller, Pearlestein, & Howard, 2001). Based on structured diagnostic interviews administered at 12 weeks postpartum to assess for postpartum depression, the researcher-developed screening questionnaire had a positive predictive value of only 33%.

Finally, based on the results of an updated meta-analysis, the Postpartum Depression Predictors Inventory (PDPI) consists of 13 risk factors related to postpartum depression (Beck, 1998, 2002b). While the researcher suggests this checklist could be completed antenatally and postnatally to update a woman's risk status, further research is warranted to determine sensitivity and specificity such that the clinical utility may be examined.

Recently, an excellent systematic review (Austin & Lumley, 2003) that summarized these preceding antenatal screening studies was published near the completion of this chapter. Sixteen studies that provided sufficient data for the researchers to calculate specific screening properties were identified; studies that were included in Table 2-2 due to this systematic review are specifically acknowledged. The aim of the reviewed studies were either to: (1) describe the development of a screening tool, (2) assess the continuity of maternal mood across the perinatal period, (3) examine risk factors and depression after birth, or (4) identify high-risk mothers to participate in a prevention trial. Outcome assessments included the EPDS and/or standardized diagnostic psychiatric interviews. No screening instrument met Austin and Lumley’s (2003) outlined criteria for routine application in the antenatal period. In summary, the unacceptably low positive predictive values in all these studies make it difficult to recommend the use of screening tools in routine antenatal care.
Table 2-2. Antenatal Screening Studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Country</th>
<th>N</th>
<th>Screening Measure</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PPV</th>
<th>% Identified as High-Risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Leverton &amp; Elliott, 1988)</td>
<td>UK</td>
<td>188</td>
<td>Leverton Questionnaire</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>53</td>
</tr>
<tr>
<td>(Appleby et al., 1994)</td>
<td>UK</td>
<td>165</td>
<td>Antenatal Screening Questionnaire</td>
<td>Correlation with antenatal EPDS = 0.34</td>
<td>Correlation with postpartum EPDS = 0.24</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>(Hobfoll et al., 1995)*</td>
<td>US</td>
<td>252</td>
<td>Schedule for Affective Disorders and Schizophrenia</td>
<td>53</td>
<td>62</td>
<td>30</td>
<td>42</td>
</tr>
<tr>
<td>(Areias et al., 1996b)</td>
<td>Portugal</td>
<td>54</td>
<td>EPDS</td>
<td>29</td>
<td>89</td>
<td>56</td>
<td>17</td>
</tr>
<tr>
<td>(Cooper et al., 1996)</td>
<td>UK</td>
<td>6091</td>
<td>Predictive Index</td>
<td>35</td>
<td>87</td>
<td>35</td>
<td>16</td>
</tr>
<tr>
<td>(Stamp et al., 1996)</td>
<td>Australia</td>
<td>248</td>
<td>Modified Antenatal Screening Questionnaire</td>
<td>73</td>
<td>43</td>
<td>17</td>
<td>58</td>
</tr>
<tr>
<td>(Posner et al., 1997)</td>
<td>US</td>
<td>106</td>
<td>Antepartum Questionnaire</td>
<td>82</td>
<td>78</td>
<td>30</td>
<td>28</td>
</tr>
<tr>
<td>(Glasser et al., 1998)*</td>
<td>Israel</td>
<td>344</td>
<td>Beck Depression Inventory</td>
<td>68</td>
<td>74</td>
<td>44</td>
<td>35</td>
</tr>
<tr>
<td>(Nhiwatiwa et al., 1998)</td>
<td>Zimbabwe</td>
<td>500</td>
<td>Shona Symptom Questionnaire</td>
<td>82</td>
<td>66</td>
<td>46</td>
<td>19</td>
</tr>
<tr>
<td>(Buist et al., 1999)</td>
<td>Australia</td>
<td>348</td>
<td>Risk Factor Scale</td>
<td>—</td>
<td>—</td>
<td>0</td>
<td>25</td>
</tr>
<tr>
<td>(Brugha et al., 2000)</td>
<td>UK</td>
<td>1300</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>19</td>
<td>31</td>
</tr>
<tr>
<td>(Elliott et al., 2000)</td>
<td>UK</td>
<td>999</td>
<td>Vulnerability Index</td>
<td>57</td>
<td>75</td>
<td>38</td>
<td>33</td>
</tr>
<tr>
<td>(Forman et al., 2000)*</td>
<td>Denmark</td>
<td>6790</td>
<td>—</td>
<td>79</td>
<td>68</td>
<td>12</td>
<td>35</td>
</tr>
<tr>
<td>(Johnson et al., 2000)*</td>
<td>UK</td>
<td>509</td>
<td>EPDS</td>
<td>23</td>
<td>91</td>
<td>17</td>
<td>10</td>
</tr>
<tr>
<td>(Webster et al., 2000)*</td>
<td>Australia</td>
<td>2118</td>
<td>Postnatal Depression Risk Index</td>
<td>29</td>
<td>89</td>
<td>32</td>
<td>14</td>
</tr>
<tr>
<td>(Josefsson et al., 2001)*</td>
<td>Sweden</td>
<td>1558</td>
<td>EPDS</td>
<td>45</td>
<td>86</td>
<td>33</td>
<td>18</td>
</tr>
<tr>
<td>(Zlotnick et al., 2001)</td>
<td>US</td>
<td>37</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>33</td>
<td>67</td>
</tr>
</tbody>
</table>

* Calculations are based on Austin and Lumley (2003)
Screening in the Immediate Postpartum

While it has been recognized that screening for postpartum depression may be beneficial in identifying depressed mothers, the actual timing of screening procedures determines whether the intervention provided takes a secondary preventive or treatment focus. Although terminology and definitions of mood disturbances in the early postpartum period has yet to be clarified and specific criteria for maternity blues has not been well established, the predictive power of maternal mood in the immediate postpartum period (e.g., first 2 weeks postpartum) has consistently been reported to be related to postpartum depression (Beck, Reynolds, & Rutowski, 1992; Fossey, Papiernik, & Bydlowski, 1997; Hapgood, Elkind, & Wright, 1988; Yoshida, Marks et al., 1997). For example, the EPDS was used to measure depressive symptomatology in 217 UK mothers at 5 days and 6 weeks postpartum (Hannah, Adams, Lee, Glover, & Sandler, 1992). A significant positive correlation between the two EPDS scores was found ($r = 0.60, p < 0.001$) and of the 25 women who scored above 12 on the EPDS at 6 weeks, 17 (68%) had similar symptomatology in the first week postpartum (5-day EPDS score above 9). In addition, mothers scoring above 9 on the EPDS at 5 days were 8 times more likely to score above 9 at 6 weeks than those scoring below 10; a previous history of postpartum depression and an EPDS score above 12 at 5 days postpartum increased the risk of postpartum depression at 6 weeks 85-fold.

In a similar but smaller study ($N = 88$), Japanese mothers scoring above 9 on the EPDS at 5 days postpartum were 20 times more likely to be diagnosed with postpartum depression during the first 12 weeks postpartum using the Schedule for Affective Disorder and Schizophrenia (SADS)/Research Diagnostic Criteria (RDC)(Yamashita, Yoshida, Nakano, & Tashiro, 2000). Predictors of mood disturbances at 3 days and 6 weeks postpartum were also assessed in 242 Irish mothers (Lane et al., 1997). Eleven percent of mothers ($n = 24$) had EPDS scores above 12 at 3 days and at 6 weeks postpartum. While factors associated with a high EPDS score at 6 weeks postpartum included single status, unemployment, unplanned pregnancy, public status, and bottle-feeding, the strongest predictor was maternal EPDS score at 3 days.

In a population-based sample of 594 Canadian mothers who completed the EPDS at 1, 4, and 8 weeks postpartum, the 1-week EPDS was significantly correlated to the 4-week ($r = 0.72, p < 0.001$) and 8-week ($r = 0.65, p < 0.001$) EPDS; a strong relationship between maternal EPDS scores at 4 and 8 weeks was also noted ($r = 0.72, p < 0.001$)(Dennis, in press-a). Using the cut-off score of 9/10, the 1-week EPDS accurately classified 457 (85.4%) mothers at 4 weeks and 410 (82.5%) mothers at 8 weeks with or without minor/major postpartum depression symptomatology; the 1-week EPDS failed to identify 3 (6%) mothers at 4 weeks and 6 (15.7%) mothers at 8 weeks who exhibited major postpartum depression symptomatology. In comparison, using a cut-off of 12/13, the 1-week EPDS accurately classified 464 (86.7%) mothers at 4 weeks and 424 (85.3%) at 8 weeks with or without major postpartum depression symptomatology. However, the 1-week EPDS failed to detect 21 (42.9%) mothers at 4 weeks and 20 (52.6%) mothers at 8 weeks who exhibited major postpartum depression symptomatology. Mothers with a 1-week EPDS score above 9 were 30.3 times
more likely at 4 weeks (95% CI = 17.5 - 42.3) and 19.1 times more likely at 8 weeks (95% CI = 11.0 - 32.9) to exhibit postpartum depression symptomatology. Using a 1-week EPDS score above 12, mothers were 11.6 times more likely at 4 weeks (95% CI = 6.1 - 21.9) and 6.9 times more likely at 8 weeks (95% CI = 3.4 - 13.8) to exhibit major postpartum depression symptomatology. In a recent meta-analysis of 85 studies (Beck, 2002a), “maternity blues” was a significant predictive factor of postpartum depression, further confirming these preceding studies that maternal mood in the immediate postpartum period is a salient factor that warrants further investigation.

Implications for Practice, Policy, and Research

While postpartum depression is moderately prevalent with 13% of new mothers experiencing this condition (O'Hara & Swain, 1996), rates including subsyndromal cases (i.e., depression that is not severe enough to meet DSM-IV criteria but still causes considerable disability) are substantially higher (Dennis, in press-a). This is clinically notable as poorer infant-mother interactions have been reported in cases of mothers with elevated depressive symptoms but whose depression was subsyndromal (Lang et al., 1996). Although health professionals can play a significant role in the detection and management of postpartum depression, this affective condition is a hidden form of maternal morbidity, often remaining undiagnosed. Researchers have identified various maternal help-seeking barriers, including the inability to identify depression indicators, fear of stigmatization, not knowing where to obtain assistance, and cultural factors. While several of these factors are common help-seeking barriers, further research is required to determine how to effectively addresses these obstacles as they pertain specifically to postpartum depression. Furthermore, options to increase knowledge among various health professionals should be examined and may include psycho-education, referral information, and practice guidelines (Boyd, Pearson, & Blehar, 2002). While competing demand models suggest that requesting health professionals to “do more” will be a challenge, significant gains in postpartum depression detection and management will not be obtainable without their systematic participation.

To aid in the detection of postpartum depression, screening procedures have been suggested. However, for a program to be effective, screening tests are required to have good sensitivity, specificity, and positive predictive values. Unfortunately, the diagnosis of postpartum depression can generally only be achieved through the application of a standardized interview by a trained mental health professional. To assist in the assessment of maternal mood, diverse self-report questionnaires have been employed. However, some measures were developed for the use in general populations (e.g., the Beck Depression Inventory and the General Health Questionnaire) resulting in unreliable scores in postpartum samples, primarily due to the similarity between the normal changes occurring in the postpartum period and symptoms indicating depression. It is also noteworthy to remember that, in general, different self-report measures assess various dimensions of the concept ‘depression’ resulting in the detection of differing subgroups. To overcome these
conceptual and psychometric limitations, postpartum depression specific questionnaires have been created. Undoubtedly, the most widely utilized instrument to screen for postpartum depression or assess maternal mood is the Edinburgh Postnatal Depression Scale (EPDS). This 10-item self-report instrument is not only convenient and acceptable to women but also easily interpretable and readily incorporated into practice. For example, the public health department in Edmonton, Alberta has completed a feasibility project with plans to screen all new mothers through the universal well-baby clinics. In the pilot stage, the researchers found good consumer satisfaction with the EPDS and that the screening could be successfully added to the task of the public health nurse (McLennan & Offord, 2002). Similar findings were found in the Fraser Valley, British Columbia where public health nurses incorporated a screening EPDS into their 8-week immunization clinics (Dennis, 2003).

While this measure has been validated among diverse cultures resulting in varying sensitivity and specificity values, it is difficult to compare research results due to the various (1) methods of assessment, (2) cut-off criteria, and (3) timing of assessments. Although these psychometric limitations are not unique to the EPDS, the methodological explanations justify only some of the discrepancies found between the EPDS translation and validation investigations. Significant differences in proportions of high EPDS scores across different cultural contexts were noted in an international multi-site study conducted by Affonso et al. (2000) suggesting that cultural factors merit further attention. In addition, further research is required to determine if indeed the EPDS is the most appropriate screening instrument, as the Postpartum Depression Screening Scale (Beck & Gable, 2000, 2001a, 2001b) has been recently developed based on qualitative interviews. As such, a comparative analysis would be prudent.

One general problem with screening instruments is that the continuous data (i.e., scores on the instruments) obtained are dichotomized into positive and negative results at an arbitrary cut-off value and then used to calculate sensitivity and specificity (as well as positive and negative predictive values). However, with this approach important information is lost as all scores above and below the threshold are counted equally. To avoid missing mothers with or at-risk for postpartum depression, a cut-off score of 9/10 is suggested for population-based screening. While this may lead to a high number of false positives and women deemed ‘at-risk,’ preliminary research suggests that a two-stage screening process may effectively address this issue (Wickberg & Hwang, 1996a, 1997). Another potential concern with the published EPDS cut-off scores are that the recommendations are based primarily on Caucasian or homogeneous samples. While a review of the presented EPDS translation and validation studies suggests that a 9/10 cut-off would be appropriate for most populations, it is unclear if this cut-off score is valid for a heterogeneous multicultural population.

Another general difficulty in measuring the accuracy of screening instruments is related to interpreting specificity. Instruments used in some studies to detect major depression may count women with
subsyndromal depression as false positives. A true measure of specificity would count as false positives only those women who are free from any significant depressive illness but who screened positive. This more accurate approach may be appropriate as women with subsyndromal illnesses may also benefit from treatment or observation that is more careful. Women with other important and treatable conditions such as anxiety, complicated grief reactions, or bipolar disorders may also be counted as false positives, but they might well be identified by the more in-depth assessment that would presumably follow a positive screen. If management of postpartum depression is initiated on only the basis of screening positive, then women with other related illness may receive sub-optimal care.

Finally, some researchers have recommended that the presence of known risk factors for postpartum depression can be employed to determine who should be screened – a strategy of selective screening. Although intuitively appealing, most common risk factors for postpartum depression perform relatively poorly in discriminating between high- and low-risk women or those who are currently depressed or not. For example, this review and the one conducted by Austin and Lumley (2003) has shown that while many different antenatal screening instruments have been created to identify women at-risk for postpartum depression, even the most well designed studies incorporating these instruments have low positive predictive values. The exclusion of salient risk factors, such as a history of depression or personality traits, is one possible explanation for the poor sensitivity and specificity of these antenatal-screening measures. However, the inclusion of postpartum variables such as birth experiences and infant mood, while potential risk factors may also limit the sensitivity and specificity of these antenatal screening measures. The need to develop a predictive tool that is clinically useful and has acceptable sensitivity and specificity remains and it has been suggested that a broader set of risk factors will need to be included (Austin & Lumley, 2003).

While there is good evidence to support the recommendation that antenatal screening to identify high-risk mothers should not be implemented into practice until additional methodological research has been completed, the saliency of maternal mood in the immediate postpartum period also warrants further exploration as a possible time to screen postnatally. Included in this research is the need to determine which time period is most effective in identifying high-risk mothers based on diverse ethnic groups. For example, researchers have reported that among Hong Kong women the postpartum supportive practice of ‘doing the month’ may have a protective effect and suggested that the pre-eminent time to screen for postpartum depression is at 6 weeks postpartum (D. Lee et al., 1998). Conversely, interviews with Caucasian women suggest that their depression began within the first 4 weeks postpartum. In multicultural communities, this poses a serious limitation for health professionals as they consider developing a systematic postpartum screening program.

While further research is required to improve screening accuracy, the effect of screening on diverse outcomes has only been partially explored. Several studies have examined the effect of screening, compared
to usual care, on the recognition and diagnosis of postpartum depression. These studies have documented that postpartum depression is often unrecognized or under-treated by ‘usual care’ or non-systematic approaches to diagnosis and management. However, significant gaps exist in the extant literature related to other postpartum depression screening outcomes. For example, few studies have been found evaluating the effect of screening on the receipt of appropriate treatment (Schaper et al., 1994), although a recently unpublished study adds to this body of literature (Chaudron, Szilagyi, Kitzman, & Conwell, 2002). In the general depression literature, a systematic review examining screening for depression in adults (Pignone et al., 2002) found the effect of screening on treatment was variable. In particular, several studies found small, non-significant increases in the proportion of patients treated for depression (Dowrick, 1995; Linn & Yager, 1980; Williams et al., 1999) while others noted an increase in antidepressant prescribing but not referral for counselling or psychiatric care (Callahan et al., 1994) or a significant 10% increase in appropriate treatment (Wells et al., 2000). Unfortunately, these results are not directly applicable to postpartum depression and further research is warranted.

Similarly, research related to the effect of screening on postpartum depression outcomes is also limited. However, a current trend in a number of countries is the preparation of guidelines and care pathways for the detection and management of pregnant and postpartum women with mental illness (Henshaw & Elliott, 2002). For example, in Scotland care pathways have been designed by a multidisciplinary group that consists of eight minimum care standards based on the recommendations presented by the SIGN National Clinical Guideline for Scotland on Postnatal Depression and Puerperal Psychosis (Robertson & Cantwell, 2002). The pathway is initiated at the primary booking for antenatal care and continues for the first year postpartum, with midwives and health visitors leading the implementation of the care standards. While the care pathway and recording tool has been piloted for 12 months and evaluated, results from this pilot have yet to be published. According to these researchers, the pathway has ensured that women are screened antenatally for puerperal psychosis risk factors and relapse of pre-existing serious mental illness. Women are offered additional support during pregnancy and high-risk women are referred to psychiatric services for prevention management. All women are screened postnatally for early signs of puerperal psychosis and the EPDS is administered to aid detection of depression at two recommended points postnatally; these time points were not reported. Women identified with mental illness are subsequently offered interventions at the appropriate level of service provision. In Scotland, the emphasis on the detection and management of mental illness antenatally and postnatally is variable with women receiving differing standards and level of care largely dependent upon the geographical area in which they lived. According to the researchers, the care pathway ensures the delivery of a minimum evidence-based standard of care for all women. While these care pathways are promising, it is unknown whether this approach either improves the number of women receiving appropriate treatment or decreases the number of women who experience postpartum depression.
cluster randomized controlled trial is required to compare ‘usual care’ with ‘care pathways’ to determine the effect on the receipt of appropriate management and postpartum depression outcomes.

In the general depression literature, the effect of screening on clinical outcome of depression is wide-ranging. In a systematic review examining screening for depression in adults (Pignone et al., 2002), two small, older trials found significant improvements in major depression (Johnstone & Goldberg, 1976; Zung & King, 1983) while two larger, well-designed trials found moderate improvements (9%) in remission from depression in a population with variable depression diagnoses (Wells et al., 2000; Williams et al., 1999); four other studies found small or no improvements in depressive outcomes (Callahan, Dittus, & Tierney, 1996; Callahan et al., 1994; Reifler, Kessler, Bernhard, Leon, & Martin, 1996; Whooley, Stone, & Soghikian, 2000). Again, these results are not directly transferable to a postpartum depression population.

In summary, several studies have examined the effect of providing feedback on postpartum depression screening results to health professionals with the rate of detection increasing from 0% to over 35%; the effect of screening on the receipt of appropriate treatment or postpartum depression outcomes was only reported by one study (Schaper et al., 1994). In the general depression literature, the results of appropriate treatment and improved depression outcomes are equivocal. Thus, although the effect of screening on diagnosis appears robust, improvements in more distal variables such as treatment and depression outcomes are unknown. Translating the increased rates of detection with screening into improved outcomes may require that particular attention be paid to initiation and maintenance of effective preventive/treatment interventions, perhaps in the form of a quality improvement effort or other programs systematically designed to provide appropriate care. Demonstrating improvements in clinical outcomes (as measured by the proportion of women still depressed) requires large samples, as studies with smaller sample sizes may be unable to demonstrate statistically significant results despite finding clinically significant differences in recovery. Furthermore, according to Pignone et al. (2002), major depression appears more responsive to interventions with screening than minor depression. As such, well-designed trials are needed and should include a discussion as to whether the appropriate outcome measure for minor postpartum depression is the same as major postpartum depression -- a failure to demonstrate changes in the proportion of women depressed may not be a reasonable test for mothers with subsyndromal illnesses.

A range of potential strategies exist in relation to screening results and include (1) simple feedback of scores obtained from screening measures, (2) feedback provided by a health professional who has received postpartum depression training, (3) feedback that can incorporate standard or individualized treatment advice, and (4) integrated recognition and management approaches that rely on multiple system supports within the clinical setting to assure prompt, coordinated follow-up. Intensive, integrated identification and management that incorporates quality improvements in clinical practice may prove to be effective in
population-based screening programs (Pignone et al., 2002). Future research comparatively evaluating these strategies should be significantly powered to detect clinically important differences in effectiveness.

It is noteworthy that no economic analysis has addressed the question of whether a modest improvement in postpartum depression outcomes warrants the increased effort of screening and providing systematic support for management (i.e., treatment or prevention). Cost-effectiveness data from two recent trials of systematic efforts to screen for general depression and provide integrated support for treatment (Schoenbaum et al., 2001; Wells et al., 2000) suggest that such programs can be implemented efficiently and produce cost-effectiveness ratios similar to those of other commonly performed preventive services, such as screening for mammography in women older than 50 years of age or treatment of mild to moderate hypertension (Pignone et al., 2002). Further research is required to determine which components of these integrated programs are most effective and to determine whether more efficient means of delivering effective care is possible.

The overarching question – whether screening and subsequent management is superior to management based on usual means of identification as ‘high-risk’– is controversial. It is unknown whether further support beyond identification improves management adherence and clinical outcomes. A recent study by Wells et al. (2000) suggests that a simple 2-question screener, when coupled with a quality improvement process, can improve outcomes over 6 to 12 months in patients with a spectrum of depressive disorders. While this has not been specifically evaluated with postpartum women, a well-designed cluster randomized controlled trial to assess community postpartum care that was redesigned to identify and manage individual needs, including postpartum depression, showed a significant decrease in maternal depression at 16-weeks postpartum (MacArthur et al., 2002). The results from this UK study suggest that screening (as part of the intervention) improved the identification of postpartum depression such that effective care could be provided. This trial should be replicated within a North America context.

The potential benefits of screening and preventing/treating postpartum depression include reduced maternal and infant morbidity, enhanced quality-of-life functioning, and improved child health outcomes; it may also decrease health service utilization (Webster et al., 2001). The potential harms of screening include (1) false positive results, (2) adverse effects of treatment, (3) negative effects and cost of treatments for women who are incorrectly identified as being depressed, and (4) potential labelling and stigmatization; there is also the question of resource implications after defining a large proportion of women as ‘at-risk’(Austin & Lumley, 2003; McLennan & Offord, 2002; Pignone et al., 2002). The trade-offs between benefits and harms are an important component in the decision of whether to screen or not. Currently, there is limited information about the harms of screening and despite a wealth of studies concerning the prevalence of postpartum depression and screening accuracy, key elements of the evidence base for screening remains insufficiently developed and a strong recommendation to implement screening procedures cannot be made. Health professionals interested in the development of postpartum depression screening programs should
proceed cautiously and observe a new US Federal Initiative that plans to screen for postpartum depression women participating in the “Health Start Program,” a comprehensive service for low income expectant and new mothers and their infants (Blehar, 2002).

Section II: Prevention of Postpartum Depression

Preventive interventions incorporate any strategy that (1) reduces the likelihood of a disease/condition affecting an individual (primary prevention), (2) interrupts or slows the progress of a disease/condition through early detection and treatment (secondary prevention), or (3) slows the progress of a disease/condition and reduces resultant disability through treatment of established disease (tertiary prevention) (Shah, 1998). These interventions can be classified into different categories depending on the target population: (1) universal measures are cost beneficial for everyone in the eligible population and target the whole population; (2) selective strategies are cost beneficial to a subgroup population who are considered to be at higher risk; and (3) indicated approaches can be applied to asymptomatic groups who have risk factors that could justify more costly and extensive interventions (Mrazek & Haggerty, 1994). Complex interactions of biopsychosocial risk factors with individual variations should be considered when planning intervention programs, as a single approach will not be applicable to all women. Standards for developing a preventive intervention have been suggested and when applied to postpartum depression should include:

- Establishing a base occurrence rate, recognizing that not all women with identified risk factors will develop postpartum depression.
- Determining the predictive accuracy of screening procedures such that vulnerable women are specifically identified.
- Being cognizant that screening procedures will exclude some women who will later develop postpartum depression.
- Devising interventions that are brief enough to be acceptable, long enough to achieve lasting benefits, intensive enough to have an effect, user friendly, and not too expensive.
- Assessing outcomes with regular monitoring and follow-up that includes a wide range of outcomes not just preventing the onset of postpartum depression.
- Recognizing that intervention non-compliance and participant attrition are major problems and that those who decline enrolment or withdraw from involvement may be those at greatest risk (Lorion, 1991).

Criteria used to assess potentially preventable conditions include the current burden of suffering (impact on the individual and on society), the manoeuvre (risks and benefits; screening accuracy; and safety, simplicity, cost, and acceptability), and intervention effectiveness (Shah, 1998). Applying these principles, postpartum depression is appropriate for preventive interventions as the long-term health consequences have
been established, there is an approximate marker of onset and a defined high-risk inception period (first 12 weeks postpartum), and women have frequent contact with health professions enabling intervention implementation (Wisner & Wheeler, 1994). Furthermore, specific knowledge about potentially modifiable risk and protective factors that influence the development of postpartum depression has been identified (as demonstrated in Chapter 1) to guide the nature of preventive strategies. However, translating risk factor research into predictive screening protocols and effective preventive interventions is challenging (Cooper et al., 1996). For this comprehensive review, preventive strategies have been classified into the following approaches: pharmacological, psychological, psychosocial, quality improvement, hormonal, and other diverse interventions. While there are a modest number of studies reporting the prevention of postpartum depression as the primary outcome, several additional investigations offer constructive data to ascertain whether limiting the influence of risk factors can decrease the incidence of postpartum depression; these studies will also be presented to provide the most comprehensive review of potential preventive interventions.

**Pharmacological Interventions**

**Antidepressant Medication**

Women who have suffered from one episode of postpartum depression are justifiably apprehensive regarding a recurrence with future births. In a naturalistic follow-up study of 20 women with initial episodes of postpartum depression who went on to have 33 more pregnancies, six mothers (30%) developed eight more incidences of postpartum depression, suggesting the risk of subsequent postpartum depression is approximately 1 in 4 (Davidson & Robertson, 1985). It has been hypothesized that administration of antidepressant medication to asymptomatic women in the immediate postpartum period may prevent recurrent episodes of postpartum depression. To determine the efficacy of prophylactic antidepressant medication, an open quasi-experimental study was conducted at a US outpatient clinic treating pregnant and postpartum women with mood disorders (Wisner & Wheeler, 1994). Twenty-three pregnant women, who had at least one previous postpartum episode that fit DSM-III-R criteria for major depression, were recruited where postpartum monitoring for recurrence of depressive symptoms (n = 8) was compared to postpartum monitoring plus antidepressant treatment with either a previously effective antidepressant medication or nortriptyline (n = 15). The first dose was given within 24 hours of birth and the recurrence of postpartum depression was monitored via psychiatric examinations for the first 12 weeks postpartum. Only one (6.7%) mother who elected postpartum monitoring plus prophylactic antidepressant medication in comparison to five (62.5%) women who elected postpartum monitoring alone suffered a recurrence (p = 0.009). However, 10 out of the 23 participants were treated with antidepressants during the current pregnancy; this included 7 (47%) in the prophylactic group and 3 (37%) in the monitoring only group. While antidepressant doses were
tapered off in the 2 weeks preceding delivery in order to recommence use again in the intervention group, the residual effect of this antenatal antidepressant use on postpartum depression reoccurrence is unknown thus limiting study conclusions.

Advancing this initial work, Wisner and colleagues conducted a double-blind, randomized controlled trial to evaluate the efficacy of nortriptyline in the prevention of recurrent postpartum depression (Wisner, Perel et al., 2001). Fifty-one non-depressed women who had at least one previous episode of postpartum depression meeting Research Diagnostic Criteria (RDC) were recruited antenatally and randomly assigned to receive either nortriptyline or a placebo in the immediate postpartum period. Each mother was assessed for 20 sequential weeks using the Hamilton Rating Scale for Depression. No significant group difference was found as 6 (23.1%) mothers who took nortriptyline prophylactically and 6 (24%) mothers who received a placebo suffered a recurrence ($p = 1.00$). Consistent with previous research, the rate of recurrence was approximately 1 in 4 women. The results from this study suggest that nortriptyline does not confer additional preventive efficacy beyond that of a placebo.

**Psychological Interventions**

**Interpersonal Psychotherapy**

Interpersonal therapy (IPT) was initially formulated as a time-limited, weekly outpatient treatment for depression provided by a trained mental health professional (Klerman & Weissman, 1993). While this method makes no assumption about aetiology, the connection between depressive symptomatology onset and interpersonal problems is used as a treatment focus. IPT as an acute treatment generally has three phases: (1) diagnosis evaluation, psychiatric/social history (including current social functioning and close relationships, their patterns, and mutual expectations), and linkage between the current interpersonal situation within one of the four interpersonal problem areas (i.e., grief, interpersonal role disputes, role transitions, or interpersonal deficits) to set the framework for treatment; (2) pursuit of strategies (defined in the IPT manual) that are specific to the chosen interpersonal problem area; and (3) encouragement to recognize and consolidate therapeutic gains and develop ways to identify and counter depressive symptoms should they arise again in the future.

To determine whether a preventive intervention based on the principles of IPT would reduce the risk of postpartum depression, two studies were found. In a US trial, 37 pregnant women receiving public assistance who had at least one risk factor for postpartum depression were randomly assigned to receive either treatment-as-usual ($n = 19$) or a group intervention (‘Survival Skills for New Moms’ consisting of four weekly 60-minute group sessions; $n = 18$) (Zlotnick et al., 2001). The majority of women in the intervention group ($n = 15$, 88%) attended three out of the four sessions and all participants completed the Beck Depression Inventory (BDI) pre and post intervention and a structured diagnostic interview (SCID) at 12...
weeks postpartum; 35 mothers completed the trial. Six (33%) women in the treatment-as-usual group developed postpartum depression in comparison to no women in the intervention group. While the results of this pilot test are promising, 50% of eligible women declined study participation rendering intervention acceptability questionable.

In a similar study, 45 pregnant US women with at least one postpartum depression risk factor (e.g., current or past history of depression, family history of treatment for psychopathology, marital problems, high levels of depressive symptomatology during pregnancy (Beck Depression Inventory score above 12), or two moderately severe life stressors) were randomly allocated to either an intervention group (five individual IPT sessions, beginning in late pregnancy and ending at approximately 4 weeks postpartum; n = 24) or a control group (standard care; n = 21) (Gorman, 2001). At 4 weeks postpartum, significantly more mothers in the control group met DSM-III-R criteria for major depression than mothers in the intervention group (25% vs. 0%, p = 0.02). However, the prophylactic effects were not maintained through 24 weeks postpartum as three (15%) mothers in the intervention group compared to four (23.5%) mothers in the control group were depressed (p = 0.40). The results from this small, underpowered trial suggest that IPT may have a limited positive preventive effect.

**Cognitive Behavioural Therapy**

Cognitive behavioural therapy (CBT) is an approach based on the notion that the way an individual perceives an event determines in part how they will respond, both affectively and behaviourally (Hollon, 1998). According to cognitive theory, dysfunctional beliefs and maladaptive information processing lie at the core of many psychiatric disorders. As such, CBT assists the individual in identifying and correcting erroneous beliefs and systematic distortions in information processing with the hopes of reducing distress and enhancing coping efforts. Only two trials have evaluated CBT as a preventive intervention for postpartum depression. In a Finnish trial, 176 pregnant women who had severe fear of childbirth were randomly allocated at 26 weeks gestation to either intensive therapy (an average of four sessions with a CBT-trained obstetrician, one session with a midwife, a recommended visit to an obstetrical ward, telephone availability between sessions, and written information regarding vaginal birth; n = 85) or conventional therapy (an average of two sessions with an obstetrician providing standard and written information regarding vaginal birth; n = 91); follow-up questionnaires, including the Beck Depression Inventory (BDI), were completed at 4 weeks before delivery date and 12 weeks postpartum (Saisto, Salmela-Aro, Nurmi, Kononen, & Halmesmaki, 2001). While birth-related concerns decreased in the intensive therapy group, no significant group differences in BDI scores were found at 12 weeks postpartum. However, postpartum depression was not the primary outcome of this study and future research evaluating CBT sessions targeting postpartum depression is needed.
In a French study, the effect of CBT targeting both the prevention and treatment of postpartum depression was evaluated (Chabrol et al., 2002); only the preventive component will be described here. Pregnant women were screened during an obstetric clinic and at-risk women (EPDS score above 8) were randomly allocated to either a control group (usual care; \( n = 128 \)) or an intervention group (\( n = 113 \)), where mothers received one individualized cognitive-behavioural session on the second or third day postpartum by a clinical therapist (which included 5 master’s level psychology students). This session comprised of three main components: (1) an educational element imparting information about the realities of parenthood, (2) a supportive element featuring empathetic listening, encouragement, and acknowledgement of feelings, and (3) a cognitive-behavioural element to “weaken the oppressive ‘shoulds’ linked to being a perfect mother.” At 4 to 6 weeks postpartum, 29 out of 97 mothers in the intervention group (29.8\%) and 55 out of 114 mothers in the control group (48.2\%) scored above 11 on the EPDS (\( \chi^2 = 7.36, p = 0.007 \)); the mean EPDS score was also significantly lower in the intervention group (\( M = 8.5, SD = 4 \)) than in the control group (\( M = 10.3, SD = 4.4; p = 0.002 \)). While these results indicate only a medium effect size (\( ES = 0.42 \)), further research is warranted.

**Psychological Debriefing**

The efficacy of psychological debriefing has been extensively debated in recent years (Arendt & Elklit, 2001) with the issues raised having ramifications beyond the field of psychological trauma (Deahl, 2000). Despite 17 years of research since the original description of “critical incident stress debriefing” (Mitchell, 1983), the role of acute interventions remains equivocal. The controversy began with articles by Bisson and Deahl (1994) and others (Raphael & Meldrum, 1995), which suggested that the beneficial effects of debriefing was next to non-existent. Consistent with this finding was a Cochrane systematic review that evaluated the effect of psychological debriefing in the prevention of post-traumatic stress disorder (Wessely, Rose, & Bisson, 2000). This meta-analysis concluded that there was no evidence debriefing prevented trauma-related symptoms and recommended ceasing compulsory debriefing of trauma victims. Two of the 11 trials included in this Cochrane review pertained to the prevention of postpartum depression. In a UK trial, 120 in-hospital primiparous women were randomized to receive either usual care (\( n = 60 \)) or a midwifery-led debriefing session before hospital discharge lasting between 30 to 120 minutes (\( n = 60 \)) (Lavender & Walkinshaw, 1998). Of the 114 women who returned the mailed Hospital Anxiety and Depression scale (HAD) at 3 weeks postpartum, significantly fewer mothers in the intervention group (\( n = 5; 8.6\%) \) exhibited depressive symptomatology in comparison to those in the control group (\( n = 31; 53.4\%) \). However, several methodological limitations existed, including premature timing of outcome assessment, poor measure of postpartum depression, atypical population (59.6\% were single women), and a ‘disappointment’ factor (e.g., dissatisfaction with group allocation) that may have increased the scores of women allocated to the control group. Conversely, in a larger and well executed Australian trial involving 1041 women who had operative
deliveries (caesarean section, \(n = 624\); use of forceps, \(n = 353\); or vacuum extraction, \(n = 64\)), in-hospital midwifery-led debriefing had a negative effect resulting in a higher rate of emotional problems (Small, Lumley, Donohue, Potter, & Waldenstrom, 2000). In particular, more women allocated to debriefing group exhibited depressive symptomatology (\(n = 81, 17\%\)) at 24 weeks postpartum than women allocated to usual postpartum care (\(n = 65, 14\%\)), although the difference was not significant (\(OR = 1.24, 95\% CI = 0.87 - 1.77\)). They were also more likely to report that depression had been a problem since delivery (\(n = 123, 28\%\) vs. \(n = 94, 22\%\)); again, the difference was not significant (\(OR = 1.37, 95\% CI = 1.00 - 1.86\)). Thus, there is strong evidence to suggest that midwifery-led debriefing after operative birth may be ineffective in reducing postpartum depression rates and the possibility that this intervention contributed to emotional health problems for some women cannot be excluded.

**Psychosocial Interventions**

**Antenatal and Postnatal Classes**

In a pioneering study, Gordon and Gordon (1960) conducted a quasi-experimental study to evaluate the effect of antenatal classes on the prevention of ‘postpartum emotional problems.’ One hundred and sixty-one pregnant US women were allocated to either a control group (standard antenatal classes; \(n = 76\)) or intervention group (standard antenatal classes plus the addition of two 40-minute sessions focusing on social and psychological adjustment; \(n = 85\)). ‘Emotional problems’ were assessed by participating obstetricians using a 4-point scale 6 to 8 weeks postpartum; ‘interjudge’ reliability was 0.85. Only 15% of mothers in the intervention group experienced emotional problems in comparison to 37% of mothers in the control group (\(\chi^2 = 7.3; p < 0.01\)). Furthermore, participants in the intervention group whose husbands attended the classes had less emotional difficulties than mothers whose husbands did not attend the classes; mothers who attended both classes had fewer difficulties than mothers who only attended one class (\(\chi^2 = 4.2, p < 0.05\)). Only half the participants completed the 24-week follow-up; one (2%) out of 46 mothers in the intervention group in comparison to 10 (28%) out of 36 mothers in the control group were experiencing emotional problems (\(\chi^2 = 9; p < 0.01\)). This study has many limitations, including non-random group allocation and a non-standardized measure of postpartum depression. Furthermore, no details were provided regarding the study groups, only that they were ‘matched by background history and were essentially the same make-up.’ Even with these weaknesses, the results suggested that the provision of realistic, solution-focused antenatal care may positively influence maternal mental health in the postpartum period and the study provided the basis for the following four antenatal class interventions.

In an Australian trial, 144 high-risk pregnant women identified using a modified antenatal screening questionnaire were randomized to receive either three midwifery-led group sessions (two antenatally and one postnatally at 6 weeks; \(n = 73\)) or standard antenatal care (\(n = 71\)) (Stamp, Williams, & Crowther, 2000).
The response rate for the mailed EPDS questionnaire at 6 and 12 weeks postpartum was 92% and 87% at 24 weeks. At 6, 12, and 24 weeks postpartum, the proportion of mothers in the intervention group with an EPDS score above 12 was 8 (13%), 7 (11%), and 9 (15%) respectively in comparison to 11 (17%), 10 (15%), and 6 (10%) mothers in the control group, indicating the intervention did not reduce postpartum depression. However, this trial has several limitations including low group attendance (31%) and a high number of women assessed as vulnerable (58%).

In a similar trial that incorporated the screening of 1300 UK women, 209 high-risk women were randomized to evaluate the effect of a structured “risk factor reducing” program titled 'Preparing for Parenthood' designed to specifically increase social support and problem-solving skills (Brugha et al., 2000). The intervention (n = 103), consisting of six structured 2-hour weekly antenatal classes and one postpartum class provided by trained nurses and occupational therapists, was compared to routine antenatal care (n = 106). Using the General Health Questionnaire Depression subscale (GHQ-D), EPDS, and the Schedules for Clinical Assessment in Neuropsychiatry (SCAN) clinical interview with a follow-up rate exceeding 90%, no significant group differences were found in the rate of postpartum depression at 12 weeks. Specifically, 16% (n = 15) of mothers in the intervention group and 19% (n = 18) of mothers in the control group had an EPDS score above 10 (OR = 0.83, 95% CI = 0.39 - 1.79); corresponding intervention and control group rates for the GHQ-D (>1) were 26% (n = 24) and 22% (n = 21) respectively. However, only 45% of the women in the intervention group attended sufficient sessions to potentially benefit.

In another UK study, a more intensive intervention titled ‘Surviving Parenthood,’ incorporating 11 monthly meetings (5 antenatally and 6 postnatally), was conducted by a psychologist and health visitor (Elliott et al., 2000). Women expecting their first or second child and designated as 'more vulnerable' were allocated to either the preventive intervention (n = 47) or a control group (n = 52) based on expected delivery date. On average, primiparous women attend 63% of the classes while group attendance by multiparous women was 36%. Significant differences in EPDS scores and a diagnosis of depression using the Present State Examination (PSE) at 12 weeks postpartum was found between primiparous but not multiparous women. The median EPDS score for primiparous women in the intervention group at 12 weeks was 3.0 (SD = 2.50) in comparison to 8.0 (SD = 4.53) in the control group (p = 0.005). For multiparous women, the median EPDS score was 6.5 (SD = 6.10) in comparison to 9.0 (SD = 6.60) in the control group. However, methodological issues, such as inadequate sample size, lack of randomization, and significant differences between participating and non-participating eligible women, render these results debatable. It is noteworthy that in all three preceding trials, a low group attendance rate was a significant limitation.
In a pilot trial, primiparous Australian women between 12 and 24 weeks gestation were screened for postpartum depression risk factors based on a researcher-developed ‘risk factor scale’ (Buist et al., 1999). A score on this scale reflected a mix of three or more risk factors related to personal/family history of depression, premenstrual syndrome, and marital/childhood difficulties. Women scoring above 7 were viewed to be ‘at-risk’ and randomly allocated to receive either standard antenatal classes (n = 21) or intervention classes (n = 23), facilitated by a midwife and another health professional which consisted of 10 structured sessions (8 antenatally and 2 postnatally) focusing on parenting and coping. The provision of support by the facilitators outside of the sessions was also available. All participants completed the Beck Depression Inventory (BDI) and EPDS at 6 and 24 weeks postpartum; 16 (70%) mothers in the intervention group and 12 (57%) mothers in the control group completed the 24-week questionnaire. No significant group differences in depressive symptomatology were found at either assessment period. In particular, mean EPDS scores at 6 and 24 weeks for mothers in the intervention group were 7.40 and 7.57 versus 9.06 and 8.09 for mothers in the control group (p > 0.05). In addition to the study limitations of a small sample size, inexplicit randomization procedures, significant group differences in baseline characteristics, and unreported intervention attendance rate, the screening questionnaire employed was not previously evaluated and no participant in either group had an EPDS score above 12 at any time, signifying the tool had poor sensitivity, specificity, and predictive power.

Intrapartum Support

As the modern obstetric era emerges, labouring women have become more isolated from the community of supporters that were once a defining feature of childbirth. Although partners and relatives are allowed to be present during delivery, a considerable number of women still experience labour without continuous support. Furthermore, obstetrical care during the past several decades has viewed labour as a high-risk situation necessitating interventions and imposed restrictions. As such, the clinical environment of childbirth may have adverse effects on psychological outcomes, including the development of postpartum depression. To test this hypothesis, two trials have been conducted evaluating the effect of doula support (i.e., labour support provided by an experienced lay woman). In a South African trial, 189 women labouring alone in a local community hospital were randomly allocated to receive either additional companionship (a minimum of 5 hours of labour support from one of three volunteer companions recruited from the community; n = 92) or usual care (n = 97) (Wolman et al., 1993). At 6 weeks postpartum, the Pitt Depression Inventory was completed during a postpartum visit; 40 mothers (21%) were lost to follow-up. Mothers receiving the supportive intervention had lower mean depression scores (M = 10.4) than mothers in the control group (M = 23.3) (p < 0.001). However, a serious study limitation was the poor measure of postpartum depression, which does not have a cut-off score for depression. This study was then continued with another 73 women enrolled and the EPDS, instead of the Pitt Depression Inventory, was administered at 1 year postpartum (Nikodem,
Nolte, Wolman, Gulmezoglu, & Hofmeyr, 1998); only 50% of mothers completed this questionnaire (64/126 mothers in the support group; 67/136 in the control group). The poorly reported study with conflicting sample size totals showed no significant group differences (intervention group $M = 11, SD = 5.31$ vs. control group $M = 11, SD = 0.60$, $p = 0.78$). In a larger US trial involving three health maintenance organization-managed hospitals, nulliparous women were randomized to receive either usual care ($n = 165$) or support from a doula ($n = 149$) (Gordon et al., 1999). Data were obtained from phone interviews conducted at 4 to 6 weeks postpartum with the results showing no significant group differences in mean postpartum depression scores. Again, a serious trial limitation was the poor measure of postpartum depression, which consisted of five items on the SF-36.

To evaluate the effectiveness of professional labour support, a randomized controlled trial with prognostic stratification by centre and parity was conducted (Hodnett et al., 2002). Thirteen Canadian and US hospitals randomized 6915 women to receive either usual care ($n = 3461$) or continuous labour support by a specially trained nurse ($n = 3454$). While the primary outcome measure was caesarean delivery rate, other outcomes included maternal mood at 6 to 8 weeks postpartum. Of the 81% of participants who returned the follow-up questionnaire, 245 (8.7%) women in the continuous labour support group had EPDS scores above 12 in comparison to 277 (10.1%) women in the usual care group ($p = 0.08$). This well conducted trial suggests that continuous labour support has no protective effect on postpartum depression.

**Supportive Interactions**

Diverse supportive interventions, including nursing home visits, home-based lay support, postpartum support groups, and self-help manuals, have been suggested to have a protective effect in the development of postpartum depression. To evaluate the effect of these supportive interventions in the prevention of postpartum depression, five trials have been conducted. In an Australian trial that targeted families where the child, for environmental reasons, was at greater risk of poor health and developmental outcomes ($N = 181$), the effect of extensive nursing home visits on diverse outcomes, including postpartum depression at 6 and 16 weeks postpartum, was evaluated (Armstrong, Fraser, Dadds, & Morris, 1999, 2000). Women were recruited in the immediate postpartum period based on self-reported vulnerability factors and randomly allocated to receive either a structured program of nurse home visiting (weekly to 6 weeks, fortnightly to 12 weeks, and monthly to 24 weeks postpartum; $n = 90$), or standard community child health services (control group; $n = 91$). Mothers who received the intervention had lower EPDS scores at 6 weeks postpartum ($M = 5.67, SD = 4.1$) than mothers in the control group ($M = 7.90, SD = 5.9$) ($p = 0.004$) with only 5.8% scoring above 12 on the EPDS in comparison to 20.7% of mothers in the control group. At the 16-week follow-up, 160 families (80 intervention, 80 control) were available for assessment and the earlier difference in EPDS scores was not maintained (intervention group $M = 5.75, SD = 5.5$; control group $M = 6.64, SD = 5.6$). While only 63% of mothers in the immediate postpartum period took the time to complete the pre-trial screening questionnaire
and significant group differences existed related to baseline characteristics, this targeted home-based intervention has several strengths including good randomization process, a power analysis, valid instruments, and low losses to follow-up.

To assess outcomes in mothers who received either the First Parent Health Visitor Scheme (FPHVS), or conventional ("generic") health visiting, retrospective data on 2113 UK families were collected during 1986-1992 as part of National Health Service (NHS) service provision (Emond et al., 2002). Prospective data were collected during 1993-1998 on 459 mothers and their children (65% acceptance rate), with outcomes assessed at 6 weeks, 1 year (93% follow-up), and 2 years (80% follow-up) via self-report questionnaires. The goal of the intervention was to “help, support, and advise mothers during their first phase of parenting” which contrasted with conventional health visiting in that it targeted primiparous mothers, emphasized empowerment, and incorporated written materials. Mothers in the intervention group were visited at home antenatally (in the third trimester), immediately after birth, at 3 weeks postpartum, and then every 5 weeks until the infant was 8 months old; approximately 20% of mothers with special difficulties continued to receive the FPHVS until 2 years of age. The results indicated that more mothers in the prospective group (FPHVS) than the retrospective comparison group (conventional health visiting) scored ‘at least’ 12 on the EPDS in the antenatal period (37% vs. 30%); these initial differences were still apparent at 6 weeks postpartum (25% vs. 19%) and at 1 year (14% vs. 9%). By 2 years, the overall number of women with depressive symptomatology had increased, but there was no meaningful difference between the two groups (17% vs. 16%). Important study limitations included the retrospective/prospective cohort design, group differences in baseline EPDS scores, and significant differences between women who agreed to participate in the study and those who refused.

Recognizing a current trend in health care, and perinatal care in particular, a recent UK trial evaluated the effect of lay support in addition to usual postpartum care provided by midwives (Morrell, Spiby, Stewart, Walters, & Morgan, 2000). Mothers were randomly allocated to receive either usual care (n = 312) or additional support (n = 311), which consisted up to 10 home visits in the first month postpartum provided by trained community postnatal support workers. At 6 weeks postpartum, there was a significant difference in EPDS scores favouring the control group (intervention group M = 7.4, SD = 5.2; control group M = 6.7, SD = 5.5; p = 0.05) and no difference at 24 weeks (intervention group M = 6.6, SD = 5.1; control group M = 6.7, SD = 5.6; p = 0.73). This well-designed trial demonstrates that unstructured support has no protective effect above the regular home visits already provided by midwives.

Finally, a randomized controlled trial with a 2 x 2 factorial design was conducted to evaluate two interventions: (1) an ‘invitation’ to a local postpartum support group run weekly by a trained midwife starting at 2 weeks postpartum and (2) a postpartum support manual, mailed at 2 weeks postpartum (Reid, Glazener, Murray, & Taylor, 2002). One thousand and four primiparous Scottish women were recruited with
83% finishing the baseline questionnaire and 71% completing the 24-week follow-up. There were no significant differences in EPDS scores between the control and intervention groups at 12 and 24 weeks either with the proportion scoring above 11 or for mean EPDS scores. The 95% CI for the difference in EPDS scores effectively excluded a change in mean score of more than 10% with either intervention. While only 40% of mothers randomized to the support group attended six or more meetings, women reported favourably on the mailed postpartum support manual. The results from this trial suggest that wide-scale provision of either support groups or self-help manuals may not be appropriate if the aim is to improve measurable mental health outcomes; further research is recommended to replicate these findings.

**Quality Improvement Interventions**

**Continuity of Care**

Based on policymakers’ suggestions that continuity of care may increase women's satisfaction, new models have been proffered, including team midwifery care. Today in the UK, midwife-managed programs of care are being implemented despite diminutive research demonstrating efficacy. To compare midwifery-managed care with shared care (i.e., care divided among midwives, hospital physicians, and general practitioners) a randomized controlled trial of 1299 pregnant women who had no adverse characteristics at booking (consent rate 82%) was conducted; postpartum depression was a psychosocial outcome (Shields, Reid, Cheyne, & Holmes, 1997). A total of 1299 women were randomly allocated to receive either midwifery-managed care (n = 648) or shared care (n = 651) with 68% returning questionnaires at 7 weeks postpartum. Women in the midwifery-managed group had significantly lower EPDS scores (M = 8.1, SD = 4.9) in comparison to mothers in the shared care group (M = 9.0, SD = 4.9; t = -2.6, 95% CI = -1.6 to -0.2, p = 0.01). However, non-significant group differences were found in relation to EPDS scores above 12 (midwifery-managed group, 71/426, 16.7% vs. shared care group, 84/362, 23.2%). It is noteworthy that women in the midwifery-managed group were significantly more likely to return their questionnaire and that a 9-item EPDS was used to determine depressive symptomatology instead of the psychometrically tested 10-item EPDS; the self-harm item was deleted.

The effect of team midwifery in the standard clinic and hospital environment was further evaluated in Australia where low-risk women in early pregnancy were randomly allocated to receive either team midwifery care (n = 495) or standard care (n = 505) (Waldenstrom, Brown, McLachlan, Forster, & Brennecke, 2000). Physicians attended most women in standard care where caregiver continuity was lacking. Based on mailed questionnaires at 8 weeks postpartum, team midwifery was associated with increased satisfaction. However, no significant group differences were found in relation to depressive symptomatology as 16% of women in the midwifery care group in comparison to 12% in the standard care group exhibited EPDS scores above 12 (p = 0.19).
Traditionally, women have been advised to attend a 6-week postpartum check-up with their primary health care provider. However, some researchers have hypothesized that postpartum care initiated earlier may either prevent or allow for the early identification and management of problems including postpartum depression. For example, in a US quasi-experimental study the effect of support initiated early by the mother and neonate’s future primary care provider (paediatrician or nurse practitioner) was evaluated (Serwint et al., 1991). Mother-neonate pairs, were randomized to either a control group (routine postpartum care including first clinic visit at 2 weeks postpartum; \( n = 122 \)) or an intervention group (nursery visit by the primary care provider at 24 to 36 hours after delivery combined with 24-hour telephone access and a provider-initiated call 2 to 3 days post-discharge to answer any further questions; \( n = 129 \)). All participants were interviewed at 8 weeks postpartum, which included the Center for Epidemiological Study of Depression Scale (CES-D). While more mothers in the intervention group made a scheduled clinic visit in the first 30 days, sought some form of care at the clinic, and tried to reach their physician by phone than mothers in the control group, no significant difference in CES-D scores was found between the two groups. In particular, mothers who received the intervention had similar CES-D scores (\( M = 11.54 \)) as mothers in the control group (\( M = 13.65 \)) (\( p = 0.11 \)) with 29% scoring above 16 on the CES-D in comparison to 39% of mothers in the control group (\( p = 0.18 \)). Study limitations include a poor randomization method and measure of postpartum depression.

In Australia, a randomized controlled trial incorporating 683 mothers was conducted to investigate whether an earlier postpartum check-up visit to a general practitioner decreased depressive symptomatology and other negative health outcomes (Gunn, Lumley, Chondros, & Young, 1998). All participants received a letter and appointment date to visit a general practitioner for a check-up: the intervention group for 1 week after hospital discharge, the control group for 6 weeks postpartum. Based on postal questionnaires (average response rate was 67.5%), the percentage of women scoring above 12 on the EPDS at 12 weeks (intervention group = 16.6% vs. control group = 13.6%; \( \chi^2 = 0.8, \ p = 0.37 \)) or 24 weeks (intervention group = 11.6% vs. control group = 12.8%; \( \chi^2 = 0.2, \ p = 0.69 \)) postpartum did not differ significantly between the two groups. The researchers of this well conducted trial concluded that to make clinically important improvements in maternal health more is required than early postpartum follow-up by general practitioners.

**Home versus Clinic Follow-Up Visit**

In addition to evaluating the timing of postpartum follow-up visits on maternal mood, the setting has also been examined. To compare the health outcomes of home versus clinic follow-up visits after early postpartum hospital discharge, 1163 medically and socially low-risk mothers with uncomplicated deliveries were randomly allocated to receive either home visits by trained nurses (\( n = 580 \)) or paediatric clinic visits by nurse practitioners or physicians (\( n = 583 \)) on the third or fourth postpartum day (Lieu et al., 2000). In
contrast with the 20-minute paediatric clinic visits, the home visits were longer (median = 70 minutes), included preventive counselling about the home environment, and involved a maternal physical examination. Diverse health outcomes, including depressive symptomatology, were assessed via telephone at 2 weeks postpartum. No significant group differences in CES-D scores (cut-off 16) were found (intervention group n = 126, 22% vs. control group n = 123, 22%). However, in this trial, only half of the women at the recruiting hospitals were eligible for participation due to stringent inclusion criteria. Furthermore, the CES-D has limited psychometric testing in the immediate postpartum period and was administered prematurely to truly evaluate the preventive effect. It is also important to note that the comparison test in this study was between a home and clinic visit after hospital discharge; a group in which mothers received no early routine follow-up was not included.

Flexible Postpartum Care

In a well-designed cluster randomized controlled trial to assess community postpartum care that was redesigned to identify and manage individual needs, 36 UK general practice clusters were randomly allocated to either an intervention (n =17) or control (n = 19) group (MacArthur et al., 2002). Midwives from the practices recruited participants and provided care in both groups. Of the 2064 participating women, 1087 (53%) were in practices randomly assigned to the intervention group (midwifery care that was extended to 12 weeks postpartum with no routine contact with general practitioners and incorporated the use of a symptom checklist and the EPDS to identify and guide the management of health needs) and 977 (47%) were in practices assigned to the control group (seven midwifery home visits to 10 to 14 days postpartum, care from health visitors thereafter with general practitioners completing routine home visits and a final 6 to 8 week check-up). Multilevel analysis accounted for possible cluster effects. In total, 801 (77%) of 1087 women in the intervention group and 702 (76%) of 977 mothers in the control group returned the 16-week postal questionnaire. Women's EPDS scores were significantly lower in the intervention group than in the control group (OR = 0.57, 95% CI = 0.43 - 0.76) with 14.4% of mothers in the intervention group scoring above 12 on the EPDS in comparison to 21.3% of mothers in the control group (p = 0.01). The numerous study strengths, including cluster design, training of midwives for intervention standardization, good randomization process, power analysis, intent-to-treat data analysis, and valid timing and measure of postpartum depression, indicate that redesigning care so that it is flexible and tailored to individual needs may help to improve women's mental health and reduce probable depression at 16 weeks postpartum.

Hormonal Interventions

Despite the fall in circulating progesterone and oestrogen in the immediate postpartum period, researchers have failed to consistently demonstrate a link between hormone levels and postpartum depression (Harris, Johns et al., 1989; Harris et al., 1996). For example, O'Hara and colleagues compared hormone
concentrations for childbearing women who became depressed versus those who did not. Frequent assays of prolactin, progesterone, estradiol, free and total estriol, and cortisol and urinary free cortisol during pregnancy and immediate postpartum revealed few differences (O'Hara, Schlechte, Lewis, & Varner, 1991). However, failure to demonstrate endocrinological evidence of hormone deficiencies does not exclude them as aetiological factors as both oestrogen and progesterone have psychoactive properties. As such, several researchers have evaluated diverse hormonal prophylaxis.

**Oestrogen Therapy**

In an open-label US study, seven women with histories of postpartum psychosis and four with histories of postpartum depression were consecutively treated with high-dose oral oestrogen immediately following delivery (Sichel, Cohen, Robertson, Ruttenberg, & Rosenbaum, 1995). None of the women had histories of non-puerperal affective disorder and all were affectively well throughout the current pregnancy. The intervention consisted of oral Premarin daily in decreasing dosages over 4 weeks. A high dose was chosen in the first few days postpartum to try and approximate term pregnancy estradiol levels before a gradual taper, designed to cushion the usual fall to follicular phase estradiol levels. Women were evaluated daily for mood and neurovegetative symptoms during the first 5 days postpartum using a DSM-III-R checklist. Follow-up was conducted at 1, 3, 6, and 12 months postpartum via clinical interview. All but one participant remained non-depressive and required no treatment with psychotropic medications during the 1-year follow-up period. The low rate of relapse in this small descriptive study suggests further research is warranted in the prophylactic ability of oral oestrogen in the immediate postpartum period among mothers at risk for a reoccurrence of postpartum affective disorders. However, it is noteworthy that research has failed to demonstrate a consistent relationship between postpartum depression and breastfeeding (which induces lower oestrogen levels) clearly challenging the claim that oestrogen therapy will be a useful preventive approach (Wisner & Stowe, 1997).

**Progesterone Therapy**

Dalton popularized the prophylactic use of progesterone for postpartum depression (Dalton, 1976, 1994). For example, in an open-label study where women who had previously experienced postpartum depression self-selected to take prophylactic progesterone treatment, a reduction from 68% to 10% was demonstrated in the reoccurrence rate (Dalton, 1985). In contrast, two double-blind randomized controlled trials of progesterone for premenstrual syndrome, which is thought by some researchers to have a similar hormonal aetiology as postpartum depression, found no significant differences between treatment and placebo groups (Freeman, Rickels, Sondheimer, & Polansky, 1995; Sampson, 1979). However, synthetic progestogens have been implicated in causing depression among women using them for contraception.
Thus, there is evidence to support the possibility that progesterone may either reduce or increase the risk of postpartum depression.

To address this question, Lawrie and colleagues conducted a double-blind randomized controlled trial to determine the effect of a long-acting progestogen contraceptive, norethisterone enanthate, administered postnatally on postpartum depression (Lawrie, Hofmeyr, De Jager et al., 1998). One hundred and eighty postpartum women using a non-hormonal method of contraception were recruited from a tertiary hospital in Johannesburg, South Africa. Women were randomly allocated within 48 hours of delivery to either a progestogen (a single dose of norethisterone enanthate 200mg [1 ml] by intramuscular injection; \( n = 90 \)) or placebo (1ml of normal saline placebo by intramuscular injection; \( n = 90 \)) group. Mothers completed the EPDS and Montgomery-Asberg Depression Rating Scale (MADRS) as part of a clinical interview at 1, 6, and 12 weeks postpartum. In comparison to the placebo group, women receiving the progestogen injection were at a significantly greater risk of developing depressive symptomatology by 6 weeks postpartum. The relative risk of scoring above 9 on the MADRS and above 11 on the EPDS for women in the intervention group was 2.56 (95% CI = 1.26 – 5.18) and 3.04 (95% CI = 1.52 – 6.08) respectively. No significant group differences were found at 12 weeks, of which the researchers hypothesized was related to the fact that only a single dose was administered. The results from this well conducted trial, incorporating good randomization and blinding methods, a power analysis, intent-to-treat data analysis, and valid measures, indicate that progestogen contraceptives should be used with caution in the postpartum period. It should also be noted that less than one-quarter of eligible women approached agreed to trial participation.

**Thyroid Function**

Research suggests that women who are positive for thyroid antibodies in pregnancy are at-risk of developing postpartum depression (Harris, Fung et al., 1989; Pop et al., 1993). To test the hypothesis that stabilizing thyroid function postnatally by administering daily thyroxine reduces the rate of occurrence and severity of associated depression, a randomized double-blind placebo-controlled trial was conducted in the UK where 100 microg of thyroxine or placebo was given daily to 446 thyroid-antibody-positive women (342 of whom were compliant) from 6 to 24 weeks postpartum (Harris et al., 2002). Maternal mood and thyroid status were assessed at 4-weekly intervals. There was no evidence that thyroxine had any effect on the occurrence of depression. This well-conducted trial provides preliminary good evidence that the higher rate of postpartum depression in thyroid-antibody-positive women is not corrected by daily administration of thyroxine. The researchers also suggested that the negative findings indicate that postpartum depression is most likely associated with known risk factors, such as negative life events, than abnormal biochemical thyroid function.
Other Interventions

Educational Strategies

Frequent contact with health professionals during pregnancy presents an ideal situation for the provision of information, with proponents of antenatal education claiming that such knowledge is a crucial factor in the maintenance of women’s health during pregnancy and their preparation for childbirth. To determine the effect of antenatal education on the prevention of postpartum depression, a randomized controlled trial was conducted in Australia (Hayes et al., 2001). Two-hundred and six primiparous women were randomized to either a control group (usual antenatal care; \( n = 103 \)) or an intervention group (\( n = 103 \)), which consisted of an educational package that included an informational booklet, a studio-quality audio-tape of one woman’s journey through postpartum depression, and an experienced midwife to guide the participant through the package. Women were given the option of receiving the intervention at either the antenatal clinic or their home between 28 to 36 weeks gestation. Depressive symptomatology was assessed using the Profile of Mood States (POMS) questionnaire, which was administered once antenatally at 12 to 28 weeks gestation and twice postnatally at 8 to 12 and 16 to 24 weeks; 188 mothers, 95 in the intervention group and 93 in the control group, completed the study protocol. No significant group difference was found on the depression subscale. Median scores for both the intervention and control groups ranged from 4.0 to 5.0 at all time periods (\( p > 0.05 \)). Serious trial limitations included the poor measure of postpartum depression and that the follow-up assessment was completed by a research assistant not blinded to group allocation. While this trial suggests that antenatal education may not prevent postpartum depression, a small descriptive Japanese study (\( N = 40 \)) found that an antenatal class provided by a psychiatrist and midwife as part of an obstetric-psychiatric liaison service that included postpartum depression information and availability of postpartum resources, may decrease the severity of postpartum depression and the time between onset of depressive symptoms and seeking professional help (Okano et al., 1998).

Relaxation with Guided Imagery

Relaxation is the state of being free from physiological and psychological tension while imagery includes all thoughts that evoke a sensory component which are not only visual but can also be in the form of auditory, motor, tactile, gustatory, and olfactory (Rees, 1995). Relaxation and imagery are often used together due to the reciprocal nature in which imagery can enhance the relaxation process and relaxation subsequently promotes image visualization. To determine the effect of relaxation with guided imagery on anxiety, depression, and self-esteem, 60 primiparous US women were recruited from a postpartum unit and randomly allocated to either a control group (4-week daily tape-recording of music for 15 minutes; \( n = 30 \)) or intervention group (4-week daily tape-recording of relaxation with guided imagery protocol for 15 minutes; \( n = 30 \)) (Rees, 1995). Using the Center for Epidemiological Studies Depression Scale (CES-D), mothers who
received the intervention had less depressive symptomatology at 4 weeks postpartum than mothers in the control group (intervention $M = 1.37$, $SD = 0.32$ vs. control $M = 1.64$, $SD = 0.53$; $t = -2.35$, $p = 0.01$). However, the inexplicit randomization and study procedures, small sample size, and weak measure of postpartum depression all make these results questionable. Furthermore, it is unknown how many women declined trial participation, rendering intervention acceptability undeterminable.
<table>
<thead>
<tr>
<th>Study</th>
<th>Design</th>
<th>Participants</th>
<th>Intervention</th>
<th>Outcome Measure</th>
<th>Results</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Antidepressant Medication</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Wisner &amp; Wheeler, 1994)</td>
<td>Quasi-experimental</td>
<td>23 US pregnant women who had at least 1 previous episode of PPD (DSM-III-R criteria) (I = 15) mothers (C = 8) mothers</td>
<td>Postpartum monitoring plus post-birth treatment with either previously used antidepressant medication or nortriptyline</td>
<td>Reoccurrence of PPD within 12 weeks Psychiatric examination</td>
<td>Significantly more women who elected monitoring alone ((62.5%)) suffered the recurrence of PPD compared to women who also received antidepressant medication ((6.7%))</td>
<td>Small sample size Non-random group allocation Participants were not blinded to treatment Potential cofounder - anti-depressant use during pregnancy by several participants Follow-up only to 12 weeks</td>
</tr>
<tr>
<td>(Wisner, Perel et al., 2001)</td>
<td>RCT</td>
<td>51 US women with a previous episode of PPD (I = 26) mothers (C = 25) mothers</td>
<td>Immediate post-birth treatment of nortriptyline</td>
<td>Reoccurrence of PPD in the first 20 weeks postpartum HRSD and RDC</td>
<td>No significant group differences were found. Of the 26 women who took nortriptyline preventively, 6 suffered a recurrence of postpartum depression while of the 25 women who took placebo, 6 suffered recurrence.</td>
<td>Small sample size</td>
</tr>
<tr>
<td><strong>Interpersonal Psychotherapy</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Zlotnick et al., 2001)</td>
<td>Pilot RCT</td>
<td>37 US pregnant women on public assistance who had at least 1 risk factor for postpartum depression (I = 18) mothers (C = 19) mothers</td>
<td>Four weekly 60-minute group sessions</td>
<td>PPD at 12 weeks BDI and structured clinical interview (SCID)</td>
<td>Significant group differences were found. 6 ((33%)) out of 18 women in control group developed PPD compared to none of the 17 women in the intervention group</td>
<td>Small sample size 50% of eligible women declined trial participation Inexplicit randomization process Atypical sample - 77% of participants were single Intervention provider unknown</td>
</tr>
<tr>
<td>(Gorman, 2001)</td>
<td>RCT</td>
<td>45 US pregnant women at-risk for PPD (I = 24) mothers (C = 21) mothers</td>
<td>Five individual sessions, beginning in late pregnancy and ending at approximately 4 weeks postpartum</td>
<td>PPD at 4 and 24 weeks Structured clinical interview for DSM-III-R (SCID)</td>
<td>At 4 weeks postpartum, significantly more women in the control group met DSM-III-R criteria for major depression than women in the intervention group ((25% \text{ vs. } 0%, p = 0.02)). Effects were not maintained through 24 weeks postpartum.</td>
<td>Small sample size Inexplicit randomization process</td>
</tr>
<tr>
<td>Study</td>
<td>Design</td>
<td>Participants</td>
<td>Intervention</td>
<td>Outcome Measure</td>
<td>Results</td>
<td>Limitations</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>-------------------------</td>
<td>---------------------------------------------------</td>
<td>------------------------------------------------------------------------------</td>
<td>----------------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>(Saisto et al., 2001)</td>
<td>RCT</td>
<td>176 Finnish pregnant women who had fear of childbirth I = 85 mothers C = 91 mothers</td>
<td>Intensive therapy (mean 3.8± 1.0 sessions with obstetrician and 1 session with midwife) vs. standard care (mean 2.0 sessions)</td>
<td>PPD at 12 weeks - BDI</td>
<td>No significant group differences were found related to PPD.</td>
<td>CBT intervention did not directly target PPD but rather fear of labour Statistical results related to PPD not reported</td>
</tr>
<tr>
<td>(Chabrol et al., 2002)</td>
<td>RCT</td>
<td>241 French women with EPDS screening score &gt;8 I = 113 mothers C = 128 mothers</td>
<td>One cognitive behavioural session before hospital discharge provided by a ‘therapist’ (included psychology graduate students)</td>
<td>PPD at 4 to 6 weeks - EPDS</td>
<td>Significant group differences were found. 29 (29.8%) mothers in the intervention group and 55 (48.2%) mothers in the control group scored above 11 on the EPDS ($\chi^2 = 7.36, p = 0.007$).</td>
<td>Weak randomization method A cut-off score of 8/9 rather than the recommended 9/10 was used to identify high-risk women and a cut-off score of 10/11 rather than the recommended 12/13 was used to assess for PPD</td>
</tr>
<tr>
<td>(Lavender &amp; Walkinshaw, 1998)</td>
<td>RCT</td>
<td>120 primiparous UK women I = 60 mothers C = 60 mothers</td>
<td>1 midwifery-led debriefing session before hospital discharge</td>
<td>PPD at 3 weeks - HAD</td>
<td>Significant group differences were found. 5 (8.6%) women in the debriefing group had depressive symptoms in comparison to 31 (53.4%) women in the control group.</td>
<td>Premature timing of outcome assessment Weak measure of PPD Atypical population -59.6% were single mothers</td>
</tr>
<tr>
<td>(Small et al., 2000)</td>
<td>RCT</td>
<td>1041 Australian women who had an operative birth I = 520 mothers C = 521 mothers</td>
<td>1 midwifery-led debriefing session before hospital discharge</td>
<td>PPD at 24 weeks - EPDS and SF-36</td>
<td>No significant group differences were found. 81 (17%) women allocated to debriefing scored as depressed at 24 weeks postpartum in comparison to 65 (14%) women allocated to usual postpartum care</td>
<td>No serious limitations</td>
</tr>
<tr>
<td>(Gordon &amp; Gordon, 1960)</td>
<td>Quasi-experimental</td>
<td>161 pregnant US women I = 85 mothers C = 76 mothers</td>
<td>Two 40-minute antenatal classes, in addition to standard prenatal classes, focusing on social and psychological adjustment</td>
<td>PPD at 6 and 24 weeks Obstetrician evaluation using a 4-point scale</td>
<td>Significant group differences were found. Only 15 % of the women in the intervention group experienced emotional upset in comparison to 37% of the women in the control group.</td>
<td>Non-random group allocation Primary outcome defined as ‘emotional upset’ Participant details lacking Unstandardized measure of PPD High attrition at 24 week follow-up</td>
</tr>
</tbody>
</table>

**Cognitive Behavioural Therapy**

**Psychological Debriefing**

**Antenatal and Postnatal Classes**

_Cindy-Lee Dennis, PhD_
<table>
<thead>
<tr>
<th>Study</th>
<th>Design</th>
<th>Participants</th>
<th>Intervention</th>
<th>Outcome Measure</th>
<th>Results</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Stamp et al., 1995)</td>
<td>RCT</td>
<td>144 ‘vulnerable’ pregnant Australian women (Modified antenatal screening questionnaire)</td>
<td>Three midwifery-led group sessions (2 antenatally and 1 postnatally at 6 weeks)</td>
<td>PPD at 6, 12, and 24 weeks - EPDS</td>
<td>No significant group differences were found.</td>
<td>High number of mothers screened vulnerable Only 31% of women attended all 3 sessions</td>
</tr>
<tr>
<td>(Brugha et al., 2000)</td>
<td>RCT</td>
<td>209 high-risk pregnant UK women (researcher developed screening tool)</td>
<td>‘Preparing for Parenthood’ - 6 structured 2-hour weekly antenatal classes and 1 postnatal class provided by trained nurse and occupational therapist</td>
<td>PPD at 12 weeks - EPDS and GHQ-O, and Clinical interview (SCAN)</td>
<td>No significant group differences were found.</td>
<td>Only 45% of women attended sufficient sessions to potentially benefit</td>
</tr>
<tr>
<td>(Elliott et al., 2000)</td>
<td>Quasi-experimental</td>
<td>99 ‘vulnerable’ pregnant UK women (Leverton Questionnaire or Crown Crisp Experiential Index)</td>
<td>“Preparation for Parenthood” - 11 monthly meetings (5 antenatally and 6 postnatally) conducted by a psychologist and health visitor</td>
<td>PPD at 12 weeks - EPDS</td>
<td>Significant group differences for primiparous women favouring the intervention group. Unsuccessful for ‘second-time’ women.</td>
<td>Non-random group allocation Significant differences between participating and non-participating eligible women Low-vulnerable mothers invited to groups to provide viable group sizes Low group attendance (36%) for multiparous mothers Study conducted between 1984 to 1985</td>
</tr>
<tr>
<td>(Buist et al., 1999)</td>
<td>Pilot RCT - Random allocation</td>
<td>44 ‘at-risk’ primiparous Australian women (researcher developed screening tool)</td>
<td>10 structured classes (8 antenatally and 2 postnatally) facilitated by a midwife and either a psychologist or nurse focusing on parenting and coping</td>
<td>PPD at 6 and 24 weeks - EPDS and BDI</td>
<td>No significant group differences were found. Mean EPDS scores at 6 and 24 weeks for women in the intervention group were 7.40 and 7.57 respectively versus 9.06 and 8.09 for women in the control group.</td>
<td>Small sample size Inexplicit randomization process Significant group differences in baseline characteristics Unreported class attendance rate Poor screening tool – at no time did any participant score above 12 on the EPDS</td>
</tr>
<tr>
<td>Study</td>
<td>Design</td>
<td>Participants</td>
<td>Intervention</td>
<td>Outcome Measure</td>
<td>Results</td>
<td>Limitations</td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>--------</td>
<td>------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------</td>
<td>-----------------</td>
<td>--------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------</td>
</tr>
<tr>
<td><strong>Intrapartum Support</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Wolman et al., 1993) (Nikodem et al., 1998)</td>
<td>RCT</td>
<td>189 nulliparous South African women (N = 92 mothers C = 97 mothers)</td>
<td>Additional companionship from 1 of 3 volunteer labour companions recruited from the community - a minimum of five hours of support</td>
<td>PPD at 6 weeks</td>
<td>Significant group differences were found at 6 weeks (mothers receiving the supportive intervention had lower mean depression scores (M = 10.4) than mothers in the control group (M = 23.3)) but not 52 weeks.</td>
<td>Poor measure of PPD at 6 weeks High attrition at 52 week follow-up Change is study protocol before completion</td>
</tr>
<tr>
<td>(Gordon et al., 1999)</td>
<td>RCT</td>
<td>314 nulliparous US women delivering in 1 of 3 HMO-managed hospitals (N = 149 mothers C = 165 mothers)</td>
<td>Provision of labour support from a trained doula</td>
<td>PPD at 4 to 6 weeks</td>
<td>No significant group differences were found.</td>
<td>High number of women in both groups excluded after randomization Weak measure of PPD Statistical results related to PPD not reported</td>
</tr>
<tr>
<td>(Hodnett et al., 2002)</td>
<td>RCT</td>
<td>6915 Canadian and US women (N = 3454 mothers C = 3461 mothers)</td>
<td>Continuous labour support by a specially trained nurse for a minimum of 80% of the time from randomization to delivery</td>
<td>PPD at 6 to 8 weeks</td>
<td>No significant group differences were found. 245 (8.7%) women in the continuous labour support group had EPDS scores above 12 in comparison to 277 (10.1%) women in the usual care group.</td>
<td>No serious methodological limitations</td>
</tr>
<tr>
<td><strong>Supportive Interactions</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Armstrong et al., 1999, 2000)</td>
<td>RCT</td>
<td>181 Australian families where the child was at a greater risk of poor health and developmental outcomes (N = 90 mothers C = 91 mothers)</td>
<td>Extensive nursing home visits (weekly to 6 weeks, fortnightly to 12 weeks, and monthly to 24 weeks)</td>
<td>PPD at 6 and 16 weeks</td>
<td>Significant group differences in EPDS mean scores were found at 6 weeks favouring the intervention group but not 16 weeks.</td>
<td>Only 63% of mothers completed pre-trial screening questionnaire Significant group differences in baseline characteristics</td>
</tr>
<tr>
<td>Study</td>
<td>Design</td>
<td>Participants</td>
<td>Intervention</td>
<td>Outcome Measure</td>
<td>Results</td>
<td>Limitations</td>
</tr>
<tr>
<td>-----------------------</td>
<td>---------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>-----------------</td>
<td>-----------------------------------------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>(Emond et al., 2002)</td>
<td>Retrospective/prospective cohort</td>
<td>Retrospective data on 2113 UK families; prospective data on 459 primiparous women and their children</td>
<td>Women in the prospective group were visited by health visitors at home antenatally (in third trimester), at the statutory primary birth visit, at 3 weeks postpartum, and then every 5 weeks until the infant was 8 months old; approximately 20% of women continued receive visits until 2 years of age.</td>
<td>PPD at 6 weeks and at 1 and 2 years EPDS</td>
<td>Significantly more women in the prospective group scored ‘at least’ 12 on the EPDS in the antenatal period (37% vs. 30%). These differences were still apparent at 6 weeks (25% vs. 19%) and at 1 year (14% vs. 9%). At 2 years there was no difference between the groups (17% vs. 16%).</td>
<td>Non-random group allocation Group differences in baseline EPDS scores Significant demographic differences between women who agreed to participate and those who refused</td>
</tr>
<tr>
<td>(Morrell et al., 2000)</td>
<td>RCT</td>
<td>623 UK women I = 311 mothers C = 312 mothers</td>
<td>Up to 10 home visits in the first postpartum month of up to 3 hours duration by a trained community postnatal support worker</td>
<td>PPD at 6 and 24 weeks EPDS</td>
<td>At 6 weeks postpartum, there was a significant difference in EPDS scores favouring the control group and no difference at 24 weeks.</td>
<td>No serious methodological limitations but theoretically weak in relation to the prevention of PPD</td>
</tr>
<tr>
<td>(Reid et al., 2002)</td>
<td>RCT</td>
<td>1004 primiparous UK women I = 753 mothers (2 different intervention groups) C = 251 mothers</td>
<td>2 interventions: (1) an invitation to a local postpartum support group run weekly by a trained midwife facilitator and (2) postpartum support manual mailed at 2 weeks postpartum</td>
<td>PPD at 12 and 24 weeks EPDS</td>
<td>There were no significant differences in EPDS scores between the control and intervention groups at 12 and 24 weeks either with the proportion scoring above 11 or for mean EPDS scores.</td>
<td>A significant number of women randomized to the support group did not attend SES bias in group attendees-more ‘middle’ than ‘working’ class mothers attended the groups Researchers question the practice of recruiting women antenatally for a postpartum intervention</td>
</tr>
<tr>
<td>(Shields et al., 1997)</td>
<td>RCT</td>
<td>1299 pregnant UK women who had no adverse characteristics I = 648 mothers C = 651 mothers</td>
<td>Total midwife care – midwife aimed to provide the majority of planned care throughout the antenatal, intrapartum, and postpartum period. Women also had an opportunity to discuss their feelings in a formal debriefing session during the last postpartum visit</td>
<td>PPD at 7 weeks EPDS</td>
<td>Women in the midwifery-managed group had significantly lower EPDS scores ($M = 8.1$, $SD = 4.9$) in comparison to mothers in the shared care group ($M = 9.0$, $SD = 4.9$) However, non-significant group differences were found in relation to EPDS scores above 12 (16.7% vs. 23.2%).</td>
<td>Inexplicit randomization process A 9-item EPDS was used instead of the psychometrically tested 10-item EPDS Participants in the intervention group were more likely to return the postal questionnaires</td>
</tr>
</tbody>
</table>

**Continuity of Care**

*Cindy-Lee Dennis, PhD*
<table>
<thead>
<tr>
<th>Study</th>
<th>Design</th>
<th>Participants</th>
<th>Intervention</th>
<th>Outcome Measure</th>
<th>Results</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Waldenstrom et al., 2000)</td>
<td>RCT</td>
<td>1000 low-risk Australian women in early pregnancy</td>
<td>Team midwifery care</td>
<td>PPD at 8 weeks</td>
<td>No significant group differences were found in relation to depressive</td>
<td>Demographic differences between questionnaire responders and non-responders</td>
</tr>
<tr>
<td></td>
<td>Random allocation</td>
<td>I = 495 mothers C = 505 mothers</td>
<td>EPDS</td>
<td></td>
<td>symptomatology as 16% of women in the team care group and 12% in the</td>
<td></td>
</tr>
<tr>
<td></td>
<td>using sealed envelopes</td>
<td></td>
<td></td>
<td></td>
<td>standard care group exhibited EPDS scores above 12.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Intent-to-treat</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Serwint et al., 1991)</td>
<td>Quasi-experimental Group allocation based on 2-week period</td>
<td>251 healthy US women I = 129 mothers C = 122 mothers</td>
<td>‘Early communication’: Routine postpartum care plus (1) visit 24-36 hours</td>
<td>PPD at 8 weeks</td>
<td>No significant group differences were found. Women who received the</td>
<td>Poor randomization method</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>after delivery from infant’s future care provider, (2) special 24-hour</td>
<td>CES-D</td>
<td>intervention had similar CES-D scores (M = 11.54) as women in the</td>
<td>Weak measure of PPD</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>telephone access to a physician via a pager for 8 weeks, (3) physician</td>
<td></td>
<td>control group (M = 13.65) with 29% scoring above 16 on the CES-D</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>initiated telephone call 2-3 days post discharge to answer questions</td>
<td></td>
<td>in comparison to 39% of women in the control group.</td>
<td></td>
</tr>
<tr>
<td>(Gunn et al., 1998)</td>
<td>RCT</td>
<td>683 healthy Australian women</td>
<td>All participants received a letter and appointment date to see a general</td>
<td>PPD at 12 and 24</td>
<td>No significant group difference between the percentages of women</td>
<td>Number of mothers randomized initially to the control and intervention</td>
</tr>
<tr>
<td></td>
<td>Telephone randomization</td>
<td></td>
<td>practitioner for a check-up: the intervention group for 1 week after hospital</td>
<td>weeks EPDS</td>
<td>scoring above 12 on the EPDS at 12 (16.6% vs. 13.6%) or 24 (11.6%</td>
<td>groups was not reported</td>
</tr>
<tr>
<td></td>
<td>Power analysis Intent-to-treat</td>
<td></td>
<td>discharge, the control group for 6 weeks postpartum.</td>
<td></td>
<td>vs. 12.8%) weeks.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>treat</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Lieu et al., 2000)</td>
<td>RCT</td>
<td>1163 medically and socially low-risk US women discharged home within 48</td>
<td>Home visit by trained nurse (60 minutes long and included a maternal physical</td>
<td>PPD at 2 weeks</td>
<td>No significant differences in CES-D scores (cut-off 16) were found at</td>
<td>Only 54% of mothers eligible for trial participation</td>
</tr>
<tr>
<td></td>
<td>Random allocation using sealed envelopes</td>
<td>48 hours I = 580 mothers C = 583 mothers</td>
<td>assessment and home preventive counselling) versus paediatric clinic visit by nurse practitioner or physician on the third or fourth day postpartum</td>
<td>CES-D</td>
<td>the 2-week interview (intervention group n = 126, 22% vs. control</td>
<td>Premature timing of outcome assessment</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>group n = 123, 21%).</td>
<td>Weak measure of PPD</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Design</td>
<td>Participants</td>
<td>Intervention</td>
<td>Outcome Measure</td>
<td>Results</td>
<td>Limitations</td>
</tr>
<tr>
<td>-------</td>
<td>--------</td>
<td>--------------</td>
<td>--------------</td>
<td>-----------------</td>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>Flexible Postpartum Care</td>
<td>(MacArthur et al., 2002)</td>
<td>RCT Computer randomization Power analysis Intent-to-treat</td>
<td>2064 UK women Only women expected to move out of the general practice were excluded I = 1087 mothers C = 977 mothers</td>
<td>Midwifery care with no routine contact with general practitioners that was extended to 12 weeks postpartum and incorporated the use of symptom checklist and the EPDS to identify health needs and guidelines for the management of these needs</td>
<td>PPD at 16 weeks EPDS</td>
<td>Women's EPDS scores were significantly lower in the intervention group than in the control group (OR = 0.57, 95% CI = 0.43 - 0.76) with 14.4% of mothers in the intervention group scoring above 12 on the EPDS in comparison to 21.3% of mothers in the control group (p = 0.01).</td>
</tr>
<tr>
<td>Oestrogen Therapy</td>
<td>(Sichel et al., 1995)</td>
<td>Open-label single group</td>
<td>7 US women with histories of postpartum psychosis and 4 women with histories of PPD</td>
<td>High-dose oral Premarin daily in decreasing dosages over 4 weeks</td>
<td>PPD at 1, 3, 6, 12 months Clinical interview</td>
<td>All but one participant remained non-depressive and required no treatment with psychotropic medications during the 1-year follow-up period.</td>
</tr>
<tr>
<td>Progesterone Therapy</td>
<td>(Lawrie, Hofmeyr, De Jager et al., 1998)</td>
<td>RCT Block randomization using a random numbers table Double blinding Power analysis Intent-to-treat</td>
<td>180 South African postpartum women using a non-hormonal method of contraception I = 90 mothers C = 90 mothers</td>
<td>Single dose of norethisterone enanthate 200mg (1 ml) by intramuscular injection at 48 hours postpartum</td>
<td>PPD at 1, 6, 12 weeks EPDS and MADRS</td>
<td>In comparison to the placebo group, women receiving the progestogen injection were at a significantly greater risk of developing depressive symptomatology by 6 weeks postpartum.</td>
</tr>
<tr>
<td>Thyroid Function</td>
<td>(Harris et al., 2002)</td>
<td>RCT Random allocation by computer-generated numbers Double blinding</td>
<td>446 UK thyroid-antibody-positive women</td>
<td>100 microg of thyroxine given daily from 6 to 24 weeks postpartum</td>
<td>PPD at 6, 12, 16, 20 and 24 weeks EPDS, MADRS, and GHQ</td>
<td>No significant group difference in rates of depression at any assessment point</td>
</tr>
<tr>
<td>Study</td>
<td>Design</td>
<td>Participants</td>
<td>Intervention</td>
<td>Outcome Measure</td>
<td>Results</td>
<td>Limitations</td>
</tr>
<tr>
<td>-------</td>
<td>--------</td>
<td>--------------</td>
<td>--------------</td>
<td>-----------------</td>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>(Hayes et al., 2001)</td>
<td>RCT Random allocation by computer-generated numbers</td>
<td>206 Australian primiparous women who were between 28 to 36 week pregnant</td>
<td>Educational package that consisted of in information booklet, audio-tape of one woman’s story of PPD, and an experienced midwife to review the package</td>
<td>PPD at 8 to 12 and 16 to 24 weeks POMS</td>
<td>No significant group difference was found on the depression subscale between the two groups. Median scores for both the intervention and control groups ranged from 4.0 to 5.0 at all time periods ($p &gt; 0.05$).</td>
<td>Weak measure of PPD Follow-up assessment completed by an unblinded research assistant</td>
</tr>
<tr>
<td>(Okano et al., 1998)</td>
<td>Descriptive study 2 groups</td>
<td>40 Japanese women who consulted a psychiatrist for postpartum depression; 18 mothers had attended a PPD class prenatally</td>
<td>One ‘Mother’s class’ in late pregnancy to provide information about the PPD, including preventive suggestions. Mothers were encouraged to obtain early psychiatric contact and resource information was provided.</td>
<td>Time of first psychiatric contact and interval between onset of illness and first interview.</td>
<td>The number of women with major PPD (SADS) was significantly higher in the non-attendant group than that in the attendant group. Mother who attended the group initiated contact with psychiatric services sooner than non-attending mothers.</td>
<td>Small sample size Retrospective design</td>
</tr>
<tr>
<td>(Rees, 1995)</td>
<td>RCT Random allocation</td>
<td>60 US primiparous women</td>
<td>Every morning for a 4-week period women followed a tape-recorded relaxation with guided imagery protocol for 15 minutes</td>
<td>PPD at 4 weeks CES-D</td>
<td>Significant group differences were found. Mothers who received the intervention had less depressive symptomatology at 4 weeks postpartum than mothers in the control group ($M = 1.37, SD = 0.32$ vs. $M = 1.64, SD = 0.53$).</td>
<td>Small sample size Inexplicit randomization and study procedures Weak measure of PPD</td>
</tr>
</tbody>
</table>

1 PPD = postpartum depression; 2 RCT = randomized controlled trial; 3 I = intervention group and C = control group.
Implications for Practice, Policy, and Research

The long-term consequences of postpartum depression suggest preventive approaches are warranted. Manipulation of a risk factor may improve the associated likelihood of developing postpartum depression through many different ways. The most obvious is to decrease the amount of exposure to a given risk factor or, alternatively, reduce the strength or mechanism of the relationship between the risk factor and postpartum depression (McLennan & Offord, 2002). However, translating risk factor research into predictive screening protocols and preventive interventions has met with limited success, as complex interactions of biopsychosocial risk factors with individual variations need to be contemplated. Over 30 studies have been examined in this review with the diverse aetiology of postpartum depression reflected in the broad range of approaches considered. Although theoretical justifications for many of these approaches have been presented, methodological limitations render intervention efficacy equivocal with scant evidence available to guide practice or policy recommendations (Table 2-4). Despite the recent upsurge of interest in this area, many questions remain unanswered resulting in a myriad of research implications.

Only two small US studies have evaluated the efficacy of prophylactic antidepressant medication (nortriptyline) and it is unknown whether the conflicting results are related to methodological limitations, inadequate drug mechanism, or intervention/approach ineffectiveness. Due to the poor quality of evidence, the effect of pharmacological interventions in the prevention of postpartum depression is unclear and this approach cannot be recommended for clinical practice. It is noteworthy that another small double-blind randomized controlled trial evaluating the effect of sertraline on the prevention of postpartum depression among 22 US mothers who had suffered one previous episode of postpartum depression was recently completed (K. Wisner et al., 2002). The results from this unpublished study suggest that sertraline may provide preventive effect. However, well-conducted randomized controlled trials are needed and should include sample sizes based on power analyses and interventions that evaluate commercially available antidepressants from diverse drug categories.

Similarly from a biological orientation, the effectiveness of hormonal interventions in the prevention of postpartum depression also needs to be rigorously examined in well-conducted randomized controlled trials. Research efforts should expand to investigations that examine women’s hormonal levels across the perinatal period from pregnancy until the resumption of normal menstrual cycles in order to delineate the potential effects of hormonal changes on depression and relapse risk. Currently, the neurochemical mechanism preventing affective relapse in high-risk women is only hypothesized and future research is necessary to clarify the role of prophylactic agents. It is noteworthy that one well-designed trial suggested synthetic progestogens increased the risk of developing depressive symptomatology (Lawrie, Hofmeyr, De Jager et al., 1998). As such, there is fair evidence to support the recommendation that long-acting progestogen contraceptives should probably not be given in the postpartum period (Lawrie, Herxheimer, & Dalton, 2000).
Several psychological interventions hold promise in the prevention of postpartum depression. Two small studies involving interpersonal psychotherapy have produced short-term positive results and provide preliminary evidence to suggest that this preventive strategy may have a positive effect on maternal mood. As such, there is a continuing need to rigorously evaluate the efficacy of this preventive strategy with future studies incorporating evaluations of maternal acceptability and the long-term effects of therapeutic gains. Also within this preventive approach is cognitive behavioural therapy, a strategy that has received limited attention in its preventive effect and further research specific to postpartum depression is recommended. Conversely, two trials evaluating the effect of psychological debriefing appear to provide beginning evidence to guide practice recommendations. While one UK study demonstrated a positive outcome, methodological weaknesses severely limit the results (Lavender & Walkinshaw, 1998). Due to the evidence obtained from a large well-designed randomized controlled trial (Small, Lumley, & Donohue, 2001), there is fair evidence to suggest that psychological debriefing in the immediate postpartum period has no protective effect on maternal mood and it is recommended that this strategy should not be implemented into practice. It is noteworthy that an unpublished trial incorporating 1745 healthy Australian mothers also found debriefing to be ineffective in reducing psychological problems in the first year after delivery (Priest et al., 2002), providing further evidence for this practice recommendation.

In general, the effectiveness of psychosocial approaches has not been satisfactorily demonstrated and well-designed studies with larger samples are required. Specifically, antenatal classes focusing on postpartum depression have repeatedly been shown to have little preventive effect. This finding may be due to methodological limitations, such as inadequate sample sizes, unrealistic effect sizes or no formal justification for sample size, large rates of participant decline and/or intervention attrition rates, or lack of adequate antenatal screening tools for identification of those “at-risk” leading to the targeting of heterogeneous “at risk” samples. Currently, there is little evidence to support the use of antenatal group interventions in heterogeneous samples of women “at-risk” for postpartum depression. However, research into structured interventions in homogeneous, symptomatic women is required; this would incorporate using an “indicated” rather than a “targeted” approach. These studies should address the previous methodological limitations and examine the efficacy for both antenatal symptoms as well as the prevention of postpartum depression. This research should be conducted before concluding that antenatal interventions have no place in the prevention of postpartum depression. It is noteworthy that poor group attendance was reported in several trials. This is a clinically significant finding noticeably demonstrating participant preference and intervention acceptance. Future research protocols would do well to take heed of this key finding.

Several studies have been found evaluating the effect of labour support, provided by both nurses and doulas. While postpartum depression was the primary outcome for only one study (Wolman et al., 1993), the trial by Hodnett et al., (2002) had sufficient power to detect the protective effect of continuous labour.
support and no significant group differences were found in the prevalence of depressive symptomatology. The results from this well-conducted trial provide good evidence to recommend that continuous labour support should not be considered as a preventive strategy for postpartum depression. Conversely, the importance of support postnatally is unknown. While one well-designed trial (Armstrong et al., 1999, 2000) with minor limitations suggested intensive nursing home visits had a beneficial effect in the first 6 weeks postpartum, the protective effect was not maintained to 16 weeks. It is interesting to note that the 16-week assessment coincided with a decrease in intervention intensity from weekly to monthly visits. A methodologically weaker cohort study also demonstrated no long-term positive effect of nursing home visits (Emond et al., 2002). Clearly, further research is warranted to examine the effectiveness of nursing home visits in the prevention of postpartum depression; the context of these visits should also be analyzed.

The importance of lay support also remains equivocal. In a well-designed randomized controlled trial, Morrell and colleagues (2002) demonstrated that the addition of home visits by a community support worker had no protective effect on postpartum depression. However, a review of the intervention activities revealed that the lay workers spent over 75% of their time providing instrumental support, such as housework and infant care, and minimal time providing emotional and appraisal (feedback) support. Methodologically strong, this trial was theoretically weak in relation to the prevention of postpartum depression. Due to the multidimensional nature of supportive interactions, the potential to positively influence health outcomes depends on the formulation of specific predictions as to which supportive functions will be the most effective for a particular type of stressor (Will & Shinar, 2000). In qualitative studies, women from diverse cultures who have suffered from postpartum depression consistently describe their feelings of loneliness, worries about maternal competence, role conflicts, and inability to cope (Chen, Wu, Tseng, Chou, & Wang, 1999; Nahas, Hillege, & Amasheh, 1999; Ritter et al., 2000; Small et al., 1994); instrumental support was not consequential. As such, it is not surprising that this trial did not have a protective effect in the prevention of postpartum depression.

Consistent with the lay support model, postpartum support groups have been hypothesized to prevent postpartum depression. However, similar to antenatal classes, postpartum group attendance rates are a clear problem as demonstrated by Reid et al. (2002). Furthermore, these researchers found a socio-economic bias in-group attendees, as “working class” mothers were less likely to attend group sessions. Theoretically, group sessions make sense due to the sharing of one’s experiences with similar others and the provision of peer support (i.e., mother-to-mother). Research suggests that this sharing interaction with a peer: (1) promotes social comparisons that normalize and validate experiences, enhances self-esteem and understanding, and reduces deviance; (2) provides reciprocal exchanges among equals that encourage a sense of belonging, worth, and control; (3) increases self-efficacy and one’s perceived ability to perform certain tasks or behaviours; and (4) enhances coping and adaptive behaviours through the discussion of problem-
solving techniques, coping strategies, and counter responses (Cohen, Underwood, & Gottlieb, 2002; Dennis, in press-b). Well-designed trials are still needed to evaluate the effect of postpartum group interventions that incorporate homogeneous samples and outcomes sensitive to peer support interventions. To assist in determining sensitive outcomes and why peer support may have a positive effect on health outcomes, the Peer Support Evaluation Inventory has been recently developed (Dennis, 2003). Throughout this review, what has become clear is that group sessions, while theoretically sound, have significant barriers to utilization. Future investigations are needed to evaluate the provision of support through different modes such as computers and the telephone.

Improving the quality of care provided to women has been another postpartum depression preventive approach. Two trials have evaluated the effect of early postpartum follow-up. While one study had several methodological limitations (Serwint et al., 1991), another well-designed trial has clearly shown no beneficial effect on maternal mental health outcomes (Gunn et al., 1998). As such, there is fair evidence to suggest that early postpartum follow-up has no preventive effect on postpartum depression and should not be recommended for clinical practice. Similarly, two large trials have evaluated the effect of midwifery-based continuity of care models on diverse maternal outcomes, including postpartum depression, and no significant group differences were found (Shields et al., 1997; Waldenstrom et al., 2000). However, results from a large, randomized controlled trial showed that flexible, individualized midwifery-based postpartum care that incorporated postpartum depression screening tools did have a positive effect in the prevention of postpartum depression (MacArthur et al., 2002). This intervention appears to be promising and a well-designed trial conducted within a North American context is needed to replicate these results.

Finally, the effect of educational strategies on prevention of postpartum depression is unknown. While an educational package informing women about postpartum depression was ineffective (Hayes et al., 2001), informing mothers about health service availability did assist mothers in seeking appropriate treatment sooner (Okano et al., 1998). This is a significant finding that warrants further investigation even though the intervention did not prevent postpartum depression, as lacking knowledge related to health service availability is an important help-seeking barrier in the detection and management of postpartum depression.

While this review clearly demonstrates that no specific approach can be strongly recommended for clinical practice, many specific research implications have been highlighted. To be most efficient in conducting this research there continues to be a need for further interdisciplinary networking among investigators with complementary research interests. For example, psychosocial intervention researchers could collaborate with health services researchers to develop and test multi-level intervention approaches embedded in service systems. To further address postpartum depression as a public health problem, the inclusion of ethnically and socio-economically diverse women in these research efforts is critical to examining the differences in depression symptoms, response rate to interventions, and health service use.
It is also necessary to present a few general comments regarding the development of preventive programs. Similar to screening initiatives, preventive interventions should be relatively simple and inexpensive. This is critical if the intervention is to be applied to a relatively large population; unless a project is feasible on a large scale, there is little utility in pursuing smaller demonstration projects. Furthermore, the risk of negative outcomes from a prevention intervention is a frequently ignored possibility. Although adverse effects are primarily thought of in treatment contexts, particularly pharmacological trials, prevention interventions also include the possibility of unfavourable events. For example, targeted prevention trials carry the risk of labelling and stigmatizing participants. Although these risks might be tolerable for those who are accurately identified and who benefit from the intervention, it may not be for those who were included in the intervention as false positives or who do not benefit from the intervention (McLennan & Offord, 2002). In addition, an increased rate of anxiety for mothers may be of real consequence, as a link between postpartum depression and child health outcomes has been demonstrated. While emphasising this may increase a mother’s willingness to accept a preventive intervention, it might also augment the mother’s level of anxiety or guilt if she perceives personal responsibility for placing her child at risk for a poor outcome, particularly if she is suffering from the cognitive distortions of depression that foster excessive guilt feelings (McLennan & Offord, 2002).

Finally, the preventive intervention should be acceptable to key stakeholders. This aspect should be considered because it is anticipated that the preventive program will be widely accepted and implemented, if it is ultimately demonstrated to be effective. Numerous stakeholders may potentially be involved in determining whether a program will obtain successful implementation. Stakeholders to consider include the general population (e.g., willingness to support the program through taxation), women (e.g., willingness to be screened and subsequently to accept the preventive program if screened positive), health professionals and administrators (e.g., willingness to devote priority to this intervention over others), and politicians (e.g., consistency of the program with their philosophy, minimal level of controversy, and potential political payoffs) (McLennan & Offord, 2002). As stakeholders can play a pivotal role in the success of a preventive program, further research should be conducted to look at this often forgotten aspect in postpartum depression research.
Table 2-4. Summary Quality of Evidence and Practice Recommendations for Preventive Interventions

<table>
<thead>
<tr>
<th>Intervention Strategy</th>
<th>Study</th>
<th>Research Design Rating¹</th>
<th>Quality Rating²</th>
<th>Classification of Recommendation³</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pharmacological</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Antidepressant Medication</strong></td>
<td>(Wisner &amp; Wheeler, 1994)</td>
<td>Quasi-experimental: II-1</td>
<td>Poor</td>
<td>I</td>
</tr>
<tr>
<td></td>
<td>(Wisner, Perel et al., 2001)</td>
<td>RCT: I</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Psychological</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Interpersonal Psychotherapy</strong></td>
<td>(Zlotnick et al., 2001)</td>
<td>Pilot RCT: I</td>
<td>Poor</td>
<td>I</td>
</tr>
<tr>
<td></td>
<td>(Gorman, 2001)</td>
<td>RCT: I</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Cognitive Behavioural Therapy</strong></td>
<td>(Saisto et al., 2001)</td>
<td>RCT: I</td>
<td>Poor</td>
<td>I</td>
</tr>
<tr>
<td></td>
<td>(Chabrol et al., 2002)</td>
<td>RCT: I</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Psychological Debriefing</strong></td>
<td>(Lavender &amp; Walkinshaw, 1998)</td>
<td>RCT: I</td>
<td>Poor</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td>(Small et al., 2000)</td>
<td>RCT: I</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Psychosocial</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Antenatal Classes</strong></td>
<td>(Gordon &amp; Gordon, 1960)</td>
<td>Quasi-experimental: II-1</td>
<td>Poor</td>
<td>I</td>
</tr>
<tr>
<td></td>
<td>(Stamp et al., 1995)</td>
<td>RCT: I</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(Brugha et al., 2000)</td>
<td>RCT: I</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(Elliott et al., 2000)</td>
<td>Quasi-experimental: II-1</td>
<td>Poor</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td>(Buist et al., 1999)</td>
<td>Pilot RCT: I</td>
<td>Poor</td>
<td></td>
</tr>
<tr>
<td><strong>Intrapartum Support</strong></td>
<td>(Wolman et al., 1993)</td>
<td>RCT: I</td>
<td>Poor</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td>(Nikodem et al., 1998)</td>
<td>RCT: I</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(Gordon et al., 1999)</td>
<td>RCT: I</td>
<td>Poor</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(Hodnett et al., 2002)</td>
<td>RCT: I</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td><strong>Supportive Interactions</strong></td>
<td>(Armstrong et al., 1999, 2000)</td>
<td>RCT: I</td>
<td>Fair</td>
<td>I</td>
</tr>
<tr>
<td></td>
<td>(Emond et al., 2002)</td>
<td>Cohort: II-2</td>
<td>Poor</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(Morrell et al., 2000)</td>
<td>RCT: I</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(Reid et al., 2002)</td>
<td>RCT: I</td>
<td>Fair</td>
<td></td>
</tr>
<tr>
<td><strong>Continuity of Care</strong></td>
<td>(Shields et al., 1997)</td>
<td>RCT: I</td>
<td>Fair</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td>(Waldenstrom et al., 2000)</td>
<td>RCT: I</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td><strong>Quality Improvement</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Early Postpartum Follow-Up by General Practitioners</strong></td>
<td>(Serwint et al., 1991)</td>
<td>Quasi-experimental: II-1</td>
<td>Poor</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td>(Gunn et al., 1998)</td>
<td>RCT: I</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td><strong>Home vs. Clinic Visit</strong></td>
<td>(Lieu et al., 2000)</td>
<td>RCT: I</td>
<td>Poor</td>
<td>I</td>
</tr>
<tr>
<td><strong>Flexible Postpartum Care</strong></td>
<td>(MacArthur et al., 2002)</td>
<td>RCT: I</td>
<td>Good</td>
<td>B</td>
</tr>
<tr>
<td>Situation</td>
<td>Reference</td>
<td>Evidence Type</td>
<td>Grade</td>
<td>Quality</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>----------------------------</td>
<td>---------------</td>
<td>-------</td>
<td>---------</td>
</tr>
<tr>
<td>Oestrogen Therapy</td>
<td>(Sichel et al., 1995)</td>
<td>Descriptive: III</td>
<td>Poor</td>
<td>I</td>
</tr>
<tr>
<td>Progesterone Therapy</td>
<td>(Lawrie et al., 1998)</td>
<td>RCT: I</td>
<td>Fair</td>
<td>D</td>
</tr>
<tr>
<td>Thyroid Function</td>
<td>(Harris et al., 2002)</td>
<td>RCT: I</td>
<td>Fair</td>
<td>I</td>
</tr>
<tr>
<td>Educational Strategies</td>
<td>(Hayes et al., 2001)</td>
<td>RCT: I</td>
<td>Poor</td>
<td>I</td>
</tr>
<tr>
<td>(Okano et al., 1998)</td>
<td>Descriptive: III</td>
<td>Poor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relaxation with Guided Imagery</td>
<td>(Rees, 1995)</td>
<td>RCT: I</td>
<td>Poor</td>
<td>I</td>
</tr>
</tbody>
</table>

1. I = evidence from randomized controlled trial(s); II-1 = evidence from controlled trial(s) without randomization; II-2 = evidence from cohort or case-control analytic studies, preferably from more than one centre or research group; II-3 = evidence from comparisons between times or places with or without the intervention, dramatic results in uncontrolled experiments could be included here; III = opinion of respected authorities, based on clinical experience, descriptive studies or reports of expert committees.

2. Good = a study (including meta-analyses or systematic reviews) that meets all design-specific criteria well; Fair = a study (including meta-analyses or systematic reviews) that does not meet (or it is not clear that it meets) at least one design-specific criterion but has no known “fatal flaw”; Poor = a study (including meta-analyses or systematic reviews) that has at least one design-specific “fatal flaw”, or an accumulation of lesser flaws to the extent that the results of the study are not deemed able to inform recommendation.

3. A = there is good evidence to recommend this approach; B = there is fair evidence to recommend this approach; C = the existing evidence is conflicting and does not allow making a recommendation for or against use of this approach, however other factors may influence decision-making; D = there is fair evidence to recommend against this approach; E = there is good evidence to recommend against this approach; I = there is insufficient evidence (in quantity and/or quality) to make a recommendation, however other factors may influence decision-making.
Section III: Treatment of Postpartum Depression

There is limited published research regarding the effectiveness of treatment for postpartum depression with approaches including pharmacological, psychological, psychosocial, hormonal, and other diverse strategies. What is unequivocal is that treating postpartum depression is a challenging undertaking that requires specific knowledge and expertise. This section is based on a literature review for each of the main treatment approaches.

Pharmacological Interventions

Antidepressant Medication

There are over 20 antidepressant medications commercially available in Canada today (Table 5) with this number expecting to increase in upcoming years (Remick, 2002). Selective Serotonin Reuptake Inhibitors (SSRIs) are newer-generation drugs that have been recommended by several researchers as the initial choice of treatment for postpartum depression (Marcus, Barry, Flynn, Tandon, & Greden, 2001; Nonacs & Cohen, 2002; Wisner, Parry, & Piontek, 2002). The literature on the use of these drugs in new mothers, especially those breastfeeding, has been rapidly expanding in recent years. However, only four studies have been found evaluating the effect of antidepressant medication specifically on postpartum depression, with three incorporating the use of SSRIs. Of these studies only one (Appleby, Warner, Whitton, & Faragher, 1997) was included in a Cochrane systematic review evaluating the effect of antidepressant drug treatment for postnatal depression (Hoffbrand, Howard, & Crawley, 2001). The purpose of this randomized controlled trial was to assess the clinical efficacy of fluoxetine, combined with at least one session of counselling in postpartum women, and included four treatment cells: fluoxetine or placebo plus one or six sessions of counselling (Appleby et al., 1997). The counselling was derived from cognitive behavioural therapy and while designed to be delivered by non-specialists after brief training, in this trial a psychologist with no previous clinical experience provided the 30 minute to 1-hour counselling sessions. Eighty-seven women who satisfied research diagnostic criteria for major (n = 51) and minor (n = 36) depression at 6 to 8 weeks postpartum participated with 61 (70%) completing the 12 weeks of treatment. Depressive symptomatology was assessed at 1, 4, and 12 weeks of treatment using the EPDS, Hamilton Rating Scale for Depression (HRSD), and a revised clinical interview. While highly significant improvements were seen in all four treatment groups, the progress in mothers receiving fluoxetine was significantly greater than in those receiving the placebo, and six sessions of counselling had a significantly greater effect than one single session. These differences were evident after 1 week, and improvement in all groups was complete after 4 weeks. The interaction between counselling and fluoxetine was not statistically significant. While it appears that both fluoxetine and cognitive-behavioural counselling are effective treatments for postpartum
depression, it should be noted that of the 188 confirmed cases of postpartum depression, 101 women refused trial participation primarily due to a reluctance to take antidepressant medication. The generalizability of the results is further limited due to differences in maternal characteristics between those who completed the trial and those who discontinued.

To determine the effectiveness of another SSRI, sertraline, in the treatment of women with depressive symptomatology that developed within 24 weeks postpartum, an 8-week, open-labelled trial was conducted (Stowe, Casarella, Landry, & Nemeroff, 1995). Twenty-six US women who fulfilled DSM-III-R criteria for major depression were treated with sertraline using an initial dose of 50mg/day, which was adjusted according to side effects and depression severity, to a maximum dose of 200mg/day. Biweekly assessments were conducted including clinical interviews (SIGH-D) and self-rated depression measures (EPDS and BDI). Twenty-one women (81%) completed the 8-week study with 20 exhibiting a salutary response as defined by a greater than 50% reduction in SIGH-D baseline scores; 14 women demonstrated complete symptom remission. While the results indicate that sertraline may be an efficacious treatment for women with postpartum depression, limitations such a small sample size, open-label single group design, homogeneous sample, and the possibility of a co-intervention (women were concurrently provided with support) render it impossible to determine whether the findings are due to the medication, psychosocial support, or both.

The effect of yet another SSRI, fluvoxamine, was evaluated in an 8-week, open-label US trial. Six women at 8 weeks postpartum identified with depressive symptomatology using the EPDS and Hamilton Rating Scale for Depression (HRSD) began fluvoxamine treatment, 50mg/day titrated to 150mg/day, and were followed with weekly clinical interviews and administration of the HRSD by a blinded assessor (Suri, Burt, Altshuler, Zuckerbrow-Miller, & Fairbanks, 2001). Repeated measures analysis of variance indicated a significant decline in depression scores over time with the greatest degree of improvement occurring between the second and third week. Like the previous study, these findings are severely limited by the small sample size, open-label single group design, and lack of a placebo control group.

Finally, an 8-week, flexible-dose, open-label study of venlafaxine (immediate release; M dose = 162.5 mg/day) was performed in a group of 15 US women who met DSM-III-R criteria for major depression with onset within the first 12 weeks postpartum (Cohen et al., 2001). Mothers were assessed at baseline and every 2 weeks across the study using the 17-item Hamilton Rating Scale for Depression (HRSD). Despite high baseline scores, treatment response was robust; 12 of the 15 women experienced remission of major depression (HRSD score below 8). These findings have the usual limitations of a small sample size, open-label single group design, and lack of a placebo control group and suggest further research evaluating the effectiveness of venlafaxine is warranted.

While these studies suggest antidepressant medication may be effective in treating postpartum depression, it is noteworthy that Hendrick and colleagues suggest women with postpartum depression may
be significantly more likely than non-postpartum women to present with anxious features, take longer to respond to pharmacotherapy, and require more antidepressant medication to obtain a therapeutic response (Hendrick, Altshuler, Strouse, & Grosser, 2000).

Antidepressant risks also exist in the postpartum period as breastfeeding provides a medium for direct infant exposure. No controlled studies of antidepressant medication during breastfeeding exist. While it is beyond the scope of this chapter to review all the different studies assessing the effects of antidepressant medication in pregnant and breastfeeding women, Table 2-5 provides a comprehensive list of studies related to foetal/infant outcomes and breastfeeding. In addition, a myriad of reviews have been published to provide further assistance (Altshuler et al., 2001; Arnon, Shechtman, & Ornoy, 2000; Chisholm & Kuller, 1997; Epperson et al., 2001; Iqbal, Sobhan, & Ryals, 2002; Llewellyn & Stowe, 1998; Marcus et al., 2001; McElhatton, 1994; Misri, Kostaras, & Kostaras, 2000; Misri & Kostaras, 2002; Newport, Hostetter, Arnold, & Stowe, 2002; Newport, Wilcox, & Stowe, 2001; Nonacs & Cohen, 1998, 2002; Stewart, 2001; Ward & Zamorski, 2002; Wisner, Gelenberg, Leonard, Zarin, & Frank, 1999; K. L. Wisner et al., 2002; Wisner, Perel, & Findling, 1996; Yoshida, Smith, & Kumar, 1999).
Table 2-5. References Related to Commercially Available Antidepressant Use during Pregnancy or Breastfeeding

<table>
<thead>
<tr>
<th>Category</th>
<th>Medication</th>
<th>Brand Name</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tricyclic Antidepressant (TCAs)</strong></td>
<td>Amitriptyline</td>
<td>Elavil</td>
<td>(Bader &amp; Newman, 1980; Breyer-Pfaff, Nill, Entenmann, &amp; Gaertner, 1995; Brixen-Rasmussen, Halgren, &amp; Jorgensen, 1982; Pittard &amp; O'Neal, 1986)</td>
</tr>
<tr>
<td></td>
<td>Clomipramine</td>
<td>Anafranil</td>
<td>(Schimmell, Katz, Shaag, Pastuszak, &amp; Koren, 1991)</td>
</tr>
<tr>
<td></td>
<td>Desipramine</td>
<td>Norpramin</td>
<td>(Sonnerr &amp; Orsulak, 1979; Stancer &amp; Reed, 1986)</td>
</tr>
<tr>
<td></td>
<td>Doxepin</td>
<td>Sinequan</td>
<td>(Frey, Scheidt, &amp; von Brenndorff, 1999; Kemp, Ilett, Booth, &amp; Hackett, 1985)</td>
</tr>
<tr>
<td></td>
<td>Imipramine</td>
<td>Tofranil</td>
<td>(Ware &amp; Devane, 1990; Weinstock, Cohen, Bailey, Blatman, &amp; Rosenbaum, 2001)</td>
</tr>
<tr>
<td></td>
<td>Protriptyline</td>
<td>Triptil</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>Trimipramine</td>
<td>Surmontil</td>
<td>—</td>
</tr>
<tr>
<td><strong>Selective Serotonin Reuptake Inhibitors (SSRIs)</strong></td>
<td>Citalopram</td>
<td>Celexa</td>
<td>(Heikkinen, Ekblad, Kero, Ekblad, &amp; Laine, 2002; Rampono et al., 2000; Schmidt, Olesen, &amp; Jensen, 2000; Spigset, Carieborg, Ohman, &amp; Norstrom, 1997)</td>
</tr>
<tr>
<td></td>
<td>Fluoxetine</td>
<td>Prozac</td>
<td>(Brent &amp; Wisner, 1998; Burch &amp; Wells, 1992; Chambers et al., 1999; Cohen et al., 2000; Goldstein, Corbin, &amp; Sundell, 1997; Hale, Shum, &amp; Grossberg, 2001; Hendrick, Stowe et al., 2001; Kristensen et al., 1999; Lester, Cucca, Andreozzi, Flanagan, &amp; Oh, 1993; Roy, Cole, Goldman, &amp; Barris, 1993; Suri et al., 2002; Taddio, Ito, &amp; Koren, 1996; Yoshida, Smith, Craggs, &amp; Kumar, 1998)</td>
</tr>
<tr>
<td></td>
<td>Paroxetine</td>
<td>Paxil</td>
<td>(Begg et al., 1999; Birnbaum et al., 1999; Hendrick, Fukuchi et al., 2001; Hendrick, Stowe, Alshuler, Hostetter, &amp; Fukuchi, 2000; Misri, Kim, Riggs, &amp; Kostaras, 2000; Ohman, Hagg, Carleborg, &amp; Spigset, 1999; Spigset, Carleborg, Norstrom, &amp; Sandlund, 1996; Stowe et al., 2000; Wisner, Findling, &amp; Perel, 2001)</td>
</tr>
<tr>
<td></td>
<td>Sertraline</td>
<td>Zoloft</td>
<td>(Alshuler et al., 1995; Birnbaum et al., 1999; Dodd, Stocky et al., 2000; Dodd, Stocky, Buist, Burrows, &amp; Norman, 2001; Epperson, Anderson, &amp; McDougle, 1997; Epperson et al., 2001; Hendrick, Fukuchi et al., 2001; Holland, 2000; Hostetter, Stowe, Strader, McLaughlin, &amp; Llewellyn, 2000; Kristensen et al., 1998; Oca &amp; Donn, 1999; Stowe et al., 1995; Stowe et al., 1997)</td>
</tr>
<tr>
<td><strong>Monoamine Oxidase Inhibitors (MAOIs)</strong></td>
<td>Moclobemide</td>
<td>Manerix</td>
<td>(Goodnick, 1994; Mayersohn &amp; Guentert, 1995; Pons et al., 1990; Rybakowski, 2001)</td>
</tr>
<tr>
<td></td>
<td>Phentolamine</td>
<td>Nardil</td>
<td>(Gracious &amp; Wisner, 1997)</td>
</tr>
<tr>
<td></td>
<td>Tranzylocpamine</td>
<td>Parmate</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>Terbutaline</td>
<td>—</td>
<td>(Boreus &amp; de Chateau, 1982; Lindberg et al., 1984; Lonnerholm &amp; Lindstrom, 1982)</td>
</tr>
<tr>
<td><strong>Others</strong></td>
<td>Amoxapine</td>
<td>Asendin</td>
<td>(Gelenberg, 1979)</td>
</tr>
<tr>
<td></td>
<td>Bupropion</td>
<td>Wellbutrin</td>
<td>(Briggs, Samson, Ambrose, &amp; Schroeder, 1993)</td>
</tr>
<tr>
<td></td>
<td>Maprotiline</td>
<td>Ludiomil</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>Mirtazapine</td>
<td>Remeron</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>Nefazodone</td>
<td>Serzone</td>
<td>(Dodd, Maguire, Burrows, &amp; Norman, 2000; Yapp et al., 2000)</td>
</tr>
<tr>
<td></td>
<td>Trazodone</td>
<td>Desyrel</td>
<td>(Verbeeck, Ross, &amp; McKenna, 1986)</td>
</tr>
<tr>
<td></td>
<td>Venlafaxine</td>
<td>Effexor</td>
<td>(Hendrick, Alshuler, Wertheimer, &amp; Dunn, 2001; Ilett et al., 1998; Ilett et al., 2002)</td>
</tr>
<tr>
<td></td>
<td>St John’s Wort</td>
<td>—</td>
<td>(Klier, Schafer, Schmid-Siegel, Lenz, &amp; Mannel, 2002)</td>
</tr>
</tbody>
</table>
Interpersonal Psychotherapy

Three studies have been found evaluating the effectiveness of interpersonal psychotherapy (IPT) on the treatment of both antepartum and postpartum depression. In a 16-week pilot study conducted with 13 pregnant US women who met DSM-III-R criteria for major depression, participants attended weekly 50-minute interpersonal sessions and completed pre and post treatment the Hamilton Rating Scale for Depression (HRSD), Beck Depression Inventory (BDI), and EPDS (Spinelli, 1997); specific intervention details were not reported. Depression ratings decreased significantly throughout the treatment program and of the 10 women available at the 12-week assessment, none reported depressive symptomatology. This preliminary work suggests a need for a larger trial to determine IPT efficacy.

In another descriptive study, US researchers used an adapted form of IPT for the treatment of postpartum depression; the modifications included an emphasis on assisting participants to resolve marital disputes and major role transitions that frequently occur in the postpartum period (Stuart & O'Hara, 1995a). Six mothers who met DSM-III-R criteria for major depression were treated for 12 weeks. Using the HRSD, BDI, and EPDS, significant changes for all measures were found post-treatment. Advancing this pilot work in a well-designed US trial, 120 postpartum women meeting DSM-IV criteria for major depression were recruited from the community and randomly assigned to either 12 weeks of IPT (n = 60) or a waiting list condition (WLC) control group (n = 60) (O'Hara et al., 2000). Follow-up data was collected via interview and self-report assessments of depressive symptomatology every 4 weeks; 99 (83%) of the 120 women completed the protocol. Mean HRSD scores of the women receiving IPT declined from 19.4 to 8.3, a significantly greater decrease than that which occurred in the WLC group (19.8 to 16.8). Similarly, mean BDI scores of the women who received IPT declined from 23.6 to 10.6 over 12 weeks, a significantly greater decrease than that which occurred in the WLC group (23.0 to 19.2). More women who received IPT recovered from their depressive episode based on HRSD scores of 6 or lower (37.5%) and BDI scores of 9 or lower (43.8%) compared with women in the WLC group (13.7% and 13.7%, respectively). Even though the outcomes assessors were not blinded to group allocation and the sample was homogeneous (e.g., Caucasian, educated, and married), these findings suggest that IPT may be an efficacious treatment for postpartum depression and represents a viable alternative to pharmacological interventions.

IPT has also been evaluated in a group modality. In an Austrian study, 17 women diagnosed with postpartum depression (DSM-IV criteria) participated in a group-based IPT intervention that consisted of two 60-minute individual sessions to explain IPT, nine weekly 90-minute group sessions, and one 60-minute individual termination session (Klier, Muzik, Rosenblum, & Lenz, 2001). Women were also provided with the telephone numbers of other group members to obtain additional support if needed. Mean score
comparisons revealed significant changes from baseline to post-treatment for both the EPDS and HRSD. At post-treatment, 10 (59%) mothers demonstrated full remission (HDRS < 9), five (29%) established partial remission (score decrease >50%), and two (12%) showed no improvement. Follow-up assessments at 24 weeks revealed a continued treatment effect. While the results indicate that group-based IPT may have positive implications for the treatment of postpartum depression, demonstrating both short-term and longer-term effects, study limitations such as the small sample size, absence of a control group, possible outcome assessment bias, and lack of intervention adherence render the results questionable. Furthermore, with the provision of telephone-based peer support, the possible effect of a co-intervention cannot be dismissed.

Cognitive Behavioural Therapy

In addition to the UK trial previously discussed that evaluated the effectiveness of fluoxetine and cognitive behavioural counselling (Appleby et al., 1997), three studies have been found incorporating a cognitive behavioural therapy (CBT) intervention in the treatment of postpartum depression. In an Australian trial, the effectiveness of CBT alone on postpartum depression was evaluated. The aims of this study were (1) to establish whether Early Childhood Nurses (ECNs) could be trained to deliver a modified CBT intervention for postpartum depression and (2) to compare the outcome of women treated with this therapy with 'ideal standard care' using non-specific counselling by ECNs with no additional training (Prendergast & Austin, 2001). Five ECNs were trained in CBT and supervised weekly. Postpartum women were recruited via regular screening by ECNs using the EPDS (score above 12) and diagnostically assessed by a clinical interview. Women with DSM-IV major depression (n =37) were then randomized to either an 'ideal standard care' group (n =20), which incorporated six weekly clinic visits or a CBT group (n =17), which consisted of six weekly one-hour home-based sessions by one of the CBT trained ECNs. Two stages of follow-up were undertaken: an interview immediately post-treatment and a postal questionnaire at 24 weeks. The training package was evaluated both by ECN completed questionnaires and analysis of taped therapy sessions. These evaluations indicated that ECNs could indeed deliver modified CBT. While there was a statistically significant difference in the EPDS scores between the two groups at baseline (CBT group M = 15.9 vs. control group M = 13.7), no group differences were found post-treatment or at the 24-week follow-up. However, there was a very high rate of recovery at initial follow-up with 70% to 80% of all participants having an EPDS score below 10. While this trial suggests that ECNs can provide a modified CBT intervention in the treatment of postpartum depression, for the majority of this sample with mild-moderate depression, perceived support from an ECN (forming an integral part of both the baseline assessment interview and control condition) appears to be as effective as modified CBT. Study limitations include a small sample size and significant group differences in baseline EPDS scores. Furthermore, 70% of control ECNs used some form of problem solving and pleasant-event scheduling strategies, providing significant similarities to the intervention.
In a French trial, pregnant women were screened with the EPDS during an obstetric clinic. Two hundred and fifty eight mothers at-risk of postpartum depression (EPDS scores above 8) were alternately assigned to either a prevention/treatment group or a control group (Chabrol et al., 2002). At 4 to 6 weeks postpartum, mothers with probable depression (EPDS scores above 10) were further assessed using the Hamilton Rating Scale for Depression (HRSD) and the Beck Depression Inventory (BDI). Participants with major depression continued in the control group (n = 30) or the intervention group (n = 18), which consisted of a cognitive-behavioural program of between five to eight 1-hour weekly home-visits (M = 6.6, SD = 1.6) that comprised of four components (i.e., supportive, educational, cognitive-behavioural, and psychodynamic). Based on HRDS, BDI, and EPDS scores, a significantly greater proportion of mothers in the intervention group recovered than mothers in the control group. In particular, recovery rates (HRSD score below 7) were 66.6% for the intervention group versus 6.6% for the control group. However, the small sample size, non-random group allocation (alternate numbers), and high initial dropout after group assignment suggest a larger well-conducted trial is required to confirm the study results.

CBT has also been evaluated using a group modality. In a pilot trial, 20 Australian women, recruited via local hospitals and maternal health centres, were eligible for participation if their postpartum depression developed within 24-weeks of delivery and they had a score above 12 on the EPDS and above 15 on the BDI (Meager & Milgrom, 1996). Consenting women were randomly allocated to either a waiting-list control group (mothers had an opportunity to participate in the treatment program once the study was completed) or an intervention group, which consisted of a ten-week, 90-minute group program, based on CBT principles that targeted postpartum depression risk factors. Women in the intervention group also exchanged telephone numbers and met outside the program. Six out of the 10 mothers in the intervention group completed the program and provided follow-up data. Following treatment, there was a significant reduction in EPDS and BDI scores, both within the intervention group and between the intervention and control groups. While the intervention resulted in a statistically significant improvement in depressive symptomatology, women were still moderately depressed. By contrast, depressive symptomatology in the control group did not change over the 10-week period. Despite the encouraging results, support from other participants outside the program was provided to women in the intervention group and eight of the 20 participants had been on antidepressant medication for at least 8 weeks before trial initiation, making it impossible to separate out the antidepressant, CBT, and peer support treatment effects. Furthermore, only six women in the intervention group and three women in the waiting list control group completed the program also raising questions about intervention acceptability. As such, further detailed studies are required to advance this pilot work.

Psychosocial Interventions
Peer Support

Detailed analyses of social support variables in predictive studies clearly suggest the following social
deficiencies significantly increase the risk of postpartum depression: (1) not having someone to talk openly
with who has shared and understood a similar problem (Brugha et al., 1998), (2) lacking an intimate
confidant or friend to converse with (Brugha et al., 1998; O'Hara, Rehm, & Campbell, 1983; Paykel, Emms,
Fletcher, & Rassaby, 1980; Romito, Saurel-Cubizolles, & Lelong, 1999), (3) not receiving support without
having to ask for it (Brugha et al., 1998), and (4) feeling socially isolated (Mills et al., 1995). Conversely,
companionship and belonging to a group of similar others has a protective effect (Cutrona, 1989). In
interviews with depressed mothers (n = 60) participating in a population-based study, women were asked for
their own explanations as to why they experienced postpartum depression; a “lack of support” and “feeling
isolated” were the most common responses (Small, Johnston, & Orr, 1997). When asked what advice they
would give to new mothers currently suffering from postpartum depression, the foremost suggestion
proffered was “find someone to talk to.” These findings support several researchers who have recommended
the provision of peer (mother-to-mother) support in a group modality for women experiencing postpartum
depression (Eastwood, 1995; Pitts, 1995). However, the results from three investigations are equivocal. In a
Canadian study, the effect of a support group was evaluated through the recruitment of mothers on the
second postpartum day who were asked to complete and return via mail a set of mood scales during the first
2 weeks postpartum (Fleming et al., 1992). Of the 1081 questionnaires distributed over a 3-year period, 781
(72%) were returned with 156 mothers scoring above the depression threshold of 35 on the “Current
Experience Scale” and either above 13 on the EPDS or 21 on the Multiple Affect Adjective Checklist.
Seventy-six mothers with depressive symptomatology (48% of all depressed mothers) and 76 non-depressed
mothers were recruited into the study. Participants were non-randomly allocated to either a support group
(eight weekly 2-hour semi-structured group sessions facilitated by two psychologists; n = 44), a ‘Group-by-
Mail’ group (to determine whether the support group effects were due to social interactions with other
women, participants in this group received scripts via mail that were adapted from the support group
sessions; n = 15), or a control group (usual postpartum care; n = 83). All groups included mothers who were
depressed and non-depressed. Participants completed the Center for Epidemiological Studies Depression
Scale (CES-D) at 6 and 20 weeks postpartum and were categorized as either depressed or non-depressed.
The ANOVA for social support versus control group at the 6-week assessment showed that ‘depressed’
women had significantly more negative feelings about themselves, their partners, and motherhood than non-
depressed mothers. At 20 weeks, although over 90% of the women in the support group reported that the
intervention was beneficial, depressed mothers showed significantly less improvement in self-image than
those in the control group and some underwent deterioration in their feelings. While the majority of
participants experienced an improvement in mood from 2 to 20 weeks postpartum regardless of group
allocation, the support group interventions did not significantly alleviate maternal depression and were
detrimental to depressed mothers' self-image. In addition to serious study limitations such as a poor measure
of postpartum depression, non-random group allocation, unequal group numbers, and a significant difference
between study groups in relation to maternal age, theoretical limitations also existed. Research suggests that
depressed individuals prefer to be with others who are depressed and that they feel worse after speaking with
non-depressed people, but not after speaking with similar others (Rosenblatt & Greenberg, 1991). As such,
the finding that depressed women felt worse after the support group meetings is not unexpected.

Recognizing this theoretical principle, a Chinese trial evaluated the effect of weekly support group
meetings for women who were all experiencing postpartum depression (Chen, Tseng, Chou, & Wang, 2000).
Mothers were recruited in-hospital on the second or third day postpartum to complete and return via mail the
Beck Depression Inventory (BDI) at 3 weeks postpartum. Eighty-five percent of mothers approached agreed
to participate (n = 941) with 414 returning the completed BDI. Sixty mothers with BDI scores above 9 were
randomized to either a support group (4 weekly semi-structured sessions facilitated by a nurse, each 1.5 to 2
hours in duration; n = 30) or a control group (usual postpartum care; n = 30). At the 4-week assessment,
mothers who attended the support sessions had significantly decreased BDI scores than mothers in the
control group. In particular, 60% (n = 18) of mothers in the control group exhibited depressive
symptomatology in comparison to only 33% (n = 9) of mothers in the support group. While this is the first
randomized controlled trial to evaluate the effectiveness of support groups, several limitations existed
including: (1) only 44% of mothers returned the screening questionnaire, (2) inexplicit randomization
method, (3) of the 115 mothers who met the inclusion criteria, only 60 were randomized and it is unknown
what happened to the other 55 potential participants, (4) unstandardized intervention as two support groups
met for five sessions instead of the scheduled four, and (5) data analysis was not based upon intent-to-treat
procedures. Thus, the positive results of this trial are questionable.

Finally, a group program for postnatally distressed Australian women and their partners was evaluated
(Morgan, Matthey, Barnett, & Richardson, 1997). The term ‘distress’ was used to indicate that no diagnostic
interview was undertaken to determine eligibility but rather women had a mixture of depressive
symptomatology based on EPDS scores above 12. The program consisted of eight weekly 2-hour sessions,
including one session for the couple, facilitated by an occupational therapist and nurse where
psychotherapeutic and cognitive-behavioural strategies were employed to assist them in dealing with their
postpartum concerns. The results from six separate groups are reported, in which 34 couples participated;
only one mother dropped out and attendance was over 90%. Seventeen mothers were simultaneously
receiving treatment by another health professional and some were on antidepressant medication. Participants
completed the EPDS and General Health Questionnaire (GHQ) during the first and last session and were
followed-up at 12 months. At program initiation, 66% of mothers had EPDS scores above 12, which
decreased to 22% at the final session, and no participant exhibited depressive symptomatology at the 12-month follow-up. While these results appear promising, the lack of a control group and the fact that over half of the mothers were receiving additional treatment for their postpartum depression, including antidepressant mediation, render the therapeutic effectiveness of these group sessions unknown.

Transcending the typical group modality, a pilot trial evaluating the effect of telephone-based peer support on postpartum depression symptomatology was conducted (Dennis, 2003). Canadian mothers who scored above 9 on the EPDS were identified through region-wide screening at the 8-week immunization clinics managed by public health nurses. Forty-two eligible and consenting mothers were randomly allocated to either a control group (standard postpartum care; n = 22) or a peer support group (standard postpartum care plus telephone-based support, initiated within 48 to 72 hours of randomization, from a mother who had previously experienced postpartum depression and had attended a 4-hour training session; n = 20). Follow-up was conducted at 4 and 8 weeks post-randomization by blinded research assistants. Significant group differences were found in probable major depressive symptomatology (EPDS score above 12) at the 4 and 8-week assessments. Specifically, at the 4-week assessment 40.9% of mothers in the control group scored above 12 on the EPDS in comparison to only 10% in the peer support group. Similar findings were found at the 8-week assessment where 52.4% of mothers in the control group continued to score above 12 on the EPDS in comparison to only 10% in the peer support group. A significant mean difference was found at the 4-week assessment between mothers in the control (M = 12.1; SD = 4.6) and peer support (M = 8.5; SD = 3.7) (t = 2.8, p = 0.008) groups. Comparable group differences were found at the 8-week assessment (t = 2.9, p = 0.006). These preliminary results suggest that telephone-based peer support may be an effective intervention and a larger randomized controlled trial will soon be underway.

Partner Support

In a Canadian trial to determine the impact of partner support in the treatment of mothers suffering from postpartum depression, women who met the DSM-IV criteria for major depressive disorder with postpartum onset were randomly allocated to either a control group (7 psycho-educational visits with a psychiatrist; n = 13) or an intervention group (7 psycho-educational visits with a psychiatrist during which the mother’s partner participated in 4 of the sessions; n = 16) (Misri, Kostaras, Fox, & Kostaras, 2000). All women were administered a set of questionnaires that included the EPDS and underwent a clinical assessment using the Mini International Neuropsychiatric Instrument (MINI) during visits one and seven. Immediately post-intervention there were no significant differences in mean EPDS scores between the intervention (M =11.4, SD = 6.2) and control (M = 14.6, SD = 7.2; p = 0.20) groups. However, at the 4-week follow-up, significant group differences were found favouring the intervention group (M = 8.6, SD = 5.2 vs. M = 14.7, SD = 7.2; p = 0.013). Study limitations included a small sample, inexplicit randomization procedures, and significant group difference in baseline characteristics: partners of the women in the intervention group a significantly
higher level of dyadic adjustment, suggesting they had a more positive appraisal of the marriage than did their control-group counterparts. This is a serious limitation considering that the intervention is the inclusion of partner support in the psycho-educational visits. Despite these considerable limitations, the initial results from this trial suggest partner support may have a measurable effect on women experiencing postpartum depression and warrants further investigation.

Non-Directive Counselling

The importance of non-directive counselling, sometimes called ‘listening visits,’ has been highlighted in the literature (Clement, 1995; Gerrard et al., 1993; Holden, 1987). To determine the effectiveness of these ‘listening visits,’ 55 UK women identified as depressed, through community-based EPDS screening at 6 weeks postpartum and a home psychiatric interview at 13 weeks, were randomized to either a control group (routine primary care) or a counselling group (eight weekly counselling visits by health visitors who received minimal training in non-directive counselling methods) (Holden, Sagovsky, & Cox, 1989). Fifty of the 55 participants completed the trial, 26 in the counselling group and 24 in the control group. After a mean time interval of 13 weeks, mothers were re-administered the standardised psychiatric interview and EPDS at home by a psychiatrist blinded to group allocation. According to RDC criteria, 18 (69%) women in the counselling group had fully recovered in comparison to only 9 (38%) women in the control group. When women in the intervention group were asked if they had received any help for their depression, 23 (88%) women responded that talking to their health visitor had been the most important recovery factor. However, one third of the counselled women did not recover despite the intervention. Of this sub-group, two had a long history of depression, another had postpartum depression previously, and a further two had a family history of depression, signifying postpartum depression occurring in the context of a continuum of psychiatric disturbances may be less likely to respond to a psychosocial intervention. It is also noteworthy that three women in each group were considered to have taken antidepressant medication at a therapeutic level. Even with the limitations of a small sample size and the possible antidepressant co-intervention, the trial results suggest that counselling by health visitors may be valuable in managing postpartum depression.

Extending the findings of Holden et al. (1989), Wickberg and Hwang (1996) conducted a population-based trial to evaluate the effect of counselling among Swedish women. Mothers participated in a two-stage screening procedure completing the EPDS at 8 and 12 weeks postpartum. Women who scored above 11 on both screening occasions were interviewed at home by a clinical psychologist, blinded to EPDS scores, at 13 weeks postpartum using the Montgomery-Asberg Depression Rating Scale (MADRS). Women who were identified as depressed according to DSM-III-R criteria were randomly allocated to receive either routine primary care (n = 16) or counselling (n = 15), which consisted of 6 weekly 1-hour counselling sessions provided in the home or clinic by a nurse who received brief training in non-directive counselling methods. Twelve (80%) women who received counselling were fully recovered after the intervention in comparison to
4 (25%) mothers in the control group. The findings of this well-conducted trial are tempered by the small sample size.

**Hormonal Interventions**

**Oestrogen Therapy**

To evaluate the effect of oestrogen on postpartum depression, a double-blind, placebo-controlled trial was conducted. Sixty-one women with major depression, which began within 12 weeks postpartum and persisted up to 18 months, were randomly allocated to either an active treatment (12 weeks of transdermal 17beta-oestradiol 200 micrograms daily alone, then 12 weeks with added cyclical dydrogesterone 10mg daily for 12 days each month; n = 34) or a placebo (placebo patches and tablets according to the same regimen; n = 27) group (Gregoire, Kumar, Everitt, Henderson, & Studd, 1996). All mothers were assessed monthly using the EPDS and a clinical psychiatric interview (Schedule for Affective Disorders and Schizophrenia). On EPDS baseline assessments, women in both groups were severely depressed (intervention group M = 21.8, SD = 3.0 vs. placebo group M = 21.3, SD = 2.9) and 47% (n = 16) of women in the intervention group and 37% (n = 10) in the control group were taking antidepressant medication. During the first 4 weeks of therapy, mothers receiving oestrogen (M = 13.3, SD = 5.7) improved significantly more than mothers in the placebo group (M = 16.5, SD = 5.3). Mothers receiving the placebo also improved over time but on average, their scores did not fall below the screening threshold for major depression for at least 16 weeks. The estimated overall treatment effect of oestrogen on the EPDS was 4.38 points (95% CI = 1.89 - 6.87). No other factors (e.g., age, psychiatric, obstetric and gynaecological history, severity and duration of current episode of depression, and concurrent antidepressant medication) influenced the response to oestrogen. This study demonstrates that transdermal oestrogen may be an effective treatment option for postpartum depression. However, further research is required to establish the minimum effective dose and duration of treatment as well as the antidepressant mechanism of oestrogen. The appropriateness of transdermal oestrogen also needs to be assessed in less severely depressed women.

Building upon a previous case study (Ahokas, Kaukoranta, & Aito, 1999), Ahokas and colleagues performed an open-label study of physiologic 17beta-oestradiol to further evaluate the treatment effect of estradiol (Ahokas et al., 2001). Twenty-three Finnish women fulfilling ICD-10 criteria for major depression with postpartum onset were consecutively recruited from a psychiatric emergency unit. Serum estradiol concentrations were measured at baseline and weekly during the 8-weeks of treatment with sublingual 17beta-estradiol; the treatment effect was assessed using the Montgomery-Asberg Depression Rating Scale (MADRS). At baseline, all women were severely depressed (MADRS total score M = 40.7; range, 35-45) and had a low serum estradiol concentration (M = 79.8 pmol/L; range, 23-140 pmol/L); in 16 (70%) mothers, the concentration was lower than the threshold value for gonadal failure. During the first week of treatment,
depressive symptoms diminished significantly resulting in a mean MADRS score of 11.0 ($Z = -4.20, p < 0.001$) and serum estradiol concentrations approached those of the follicular phase ($M = 342 +/- 141$ pmol/L). At the end of the second treatment week, the MADRS scores were compatible with clinical recovery in 19 (82.6%) mothers. This initial study illustrates that depressive symptomatology may be rapidly reduced in women who have documented estradiol deficiency through the treatment with 17beta-estradiol. Further research is required to determine the significance of estradiol in the pathophysiology of postpartum depression.

*Other Interventions*

**Relaxation/Massage Therapy**

Massage and relaxation therapies have been shown to decrease anxiety and elevate mood (McLean & Hakstian, 1979; Reynolds & Coats, 1986). To determine the effects of massage and relaxation therapies on postpartum depression, 32 US in-hospital adolescent mothers, who were determined to be depressed based on a Beck Depression Inventory (BDI) score above 16, were recruited and randomly allocated to either a massage therapy group (30-minute massage per day on two consecutive days per week for 5 consecutive weeks; $n = 16$) or relaxation therapy group (30-minute relaxation therapy session, consisting of 15 minutes of yoga and 15-minutes of progressive muscle relaxation, on two consecutive days per week for 5 consecutive weeks for a total of 10 sessions; $n = 16$) (Field, Grizzle, Scafidi, & Schanberg, 1996). The effects of the massage and relaxation therapies were assessed pre and post treatment on the first and last day of the sessions using the Profile of Mood States 14 item depression subscale (POMS-D). Results suggest that there was no difference in pre- and post-treatment POMS-D scores in relation to relaxation therapy but a significant difference in pre and post treatment scores on day 1 and 10 in relation to massage therapy. These results signify that unlike relaxation therapy, massage therapy may have a significant immediate effect on depression scores. However, the long-term effect of these therapies is unknown resulting in questionable clinical utility. Furthermore, the inexplicit randomization and trial procedures, small sample size, lack of a true control group, and poor measure of postpartum depression render these results equivocal.

While diminishing maternal depression does not necessarily improve mother-infant interactions, direct attempts to enhance the quality of mother-infant interactions, independently of improving maternal depression, have been reported with some success. To determine whether attending regular massage classes could reduce maternal depression and also enrich the quality of mother-infant interactions, a trial was conducted involving 34 primiparous UK mothers identified as being depressed following the completion of the EPDS at 4 weeks postpartum (Onozawa, Glover, Adams, Modi, & Kumar, 2001). Participants were randomly allocated to either an intervention group (five weekly 1-hour infant massage classes and a 30-minute informal support group; $n = 19$) or a control group (five weekly informal support groups; $n = 15$).
Changes in maternal depression were assessed at the beginning and end of the trial using the EPDS. Twelve mothers in the intervention group and 13 mothers in the control group completed all sessions (73.5%). Results suggest that there was a greater improvement in EPDS scores in the intervention group that in the control group; the median EPDS score for the intervention group at the final session was 5.0 (95% CI = 8.0 – 14.2) in comparison to 10.0 (95% CI = 4.6 – 9.0) for mothers in the control group. However, it should be noted that much of the effect occurred before the classes began, possibly reflecting expectation. The small sample size, inexplicit randomization procedures, lack of intent-to-treat data analysis, and inability to distinguish the contributing aspects of the infant massage class all diminish trial validity. Furthermore, only 25 mothers completed all the sessions rendering maternal acceptability debatable.

**Bright Light Therapy**

While bright light therapy has been shown to be an effective treatment for seasonal affective disorder and non-seasonal depression, two preliminary case report studies suggest that it may also have a beneficial effect on postpartum depression. For example, Corral, Kuan, and Kostaras (2000) report the cases of two women, suffering from postpartum depression and refusing to take antidepressant medication, who consented to a 4-week trial of phototherapy by means of a 10,000-lux light box for 30 minutes a day. Baseline Hamilton Rating Scale for Depression (HRSD) scores (29-item version) for both mothers were above 27 with each showing a 75% reduction in HRSD scores at their last treatment session (scores were 11 and 12). While no adverse effects during the course of treatment were reported, it is unknown whether the treatment effect was maintained once the phototherapy ended. It should also be noted that one mother felt her poor marital relationship precipitated her depression indicating a psychosocial aetiology and that the observed improved mood may be related to the daily social interaction received during treatment.

To further explore the use of bright light therapy, a study was conducted among 16 pregnant US women with major depression (Oren et al., 2002). Treatment consisted of ultraviolet fluorescent light incorporating a 100,000-lux box for 60 minutes daily beginning within 10 minutes of awakening for at least 3 to 5 weeks; compliance was monitored through daily answering machine reports of light use. The Hamilton Rating Scale for Depression (HRSD) was administered to assess treatment effect. After 3 weeks of treatment, mean HRSD scores improved by 49% with benefits seen through the 5 weeks of treatment; there was no evidence of adverse effects. These data provide support that bright light therapy may have an antidepressant effect and while it is evident that additional research is required, this treatment option may be a viable alternative for severely depressed mothers who are not responsive to traditional approaches.

**Maternal and Infant Sleep Interventions**

Some researchers have suggested the relationship between maternal sleep deprivation and postpartum depression. To determine the efficacy of critically timed sleep deprivation in major mood disorders (MMD)
occurring during pregnancy and postpartum, nine women who met DSM-IV criteria for a MMD with onset during pregnancy or within 1 year postpartum underwent a session of either early-night sleep deprivation (ESD), in which they were sleep deprived in the early part of one night and slept from 03:00-07:00 h, or late-night sleep deprivation (LSD), in which they were deprived of sleep in the latter part of one night and slept from 21:00-01:00 h (Parry et al., 2000). After 1 week of regular sleep, mothers who relapsed were crossed-over to the alternate sleep deprivation condition. Depressive symptomatology was assessed pre and post intervention and after a night of recovery sleep (sleep 22:30-06:30 h) by trained clinicians, blinded to treatment condition, using Hamilton Rating Scale for Depression (HRSD) and Beck Depression Inventory (BDI). More participants responded to LSD (nine of 11 sessions: 82%) compared with ESD (two of six sessions: 33%) and they responded more after a night of recovery sleep (nine of 11 nights: 82%) than after a night of sleep deprivation (six of 11 nights: 55%). Pregnant women were the only responders to ESD and the only non-responders to LSD. This study has severe methodological limitations, such as the small and heterogeneous sample size, and meaningful clinical utility has not been demonstrated making the feasibility of conducting a larger study debateable.

Commonly used behavioural interventions have been shown to decrease infant sleep problems and maternal reports of depressive symptomatology. However, uncontrolled trials, small sample sizes, and short follow-up render the results equivocal (Armstrong, Van Haeringen, Dadds, & Cash, 1998; Leeson, Barbour, Romaniuk, & Warr, 1994; Rickert & Johnson, 1988). To address this issue a well-designed trial was conducted with 156 Australian mothers of infants aged 6-12 months with severe sleep problems. Participants were recruited from well child clinics to compare the effect of a behavioural sleep intervention on infant sleep problems and maternal depression (Hiscock & Wake, 2002). Mothers in the intervention group attended three private consultations with a senior paediatric trainee, held every 2 weeks at their local maternal and child health centre; sleep management plans were also tailored according to individual needs (n = 76). Mothers also received information about the development and management of sleep problems and an information sheet about normal sleep patterns. Mothers in the control group were mailed only the information sheet (n = 76). All participants completed the EPDS at 8 and 16 weeks post-randomization. At 8 weeks, more sleep problems had resolved in the intervention group than in the control group and depressive symptomatology scores decreased more in the intervention group (mean change = -3.7, 95% CI = -4.7 to -2.7) than in the control group (mean change = -2.5, 95% CI = -1.7 to -3.4) (p = 0.06). For the subgroup of mothers with baseline EPDS scores above 9, depression scores fell significantly further for mothers in the intervention group (mean change = -6.0, 95% CI = -7.5 to -4.0) than mothers in the control group (mean change = -3.7, 95% CI = -4.9 to -2.6) (p = 0.01) at 8 weeks; similar results were found at 16 weeks (p = 0.04). These findings suggest an infant sleep modification intervention may significantly decrease depressive symptomatology, especially among mothers with high depression scores.
Electroconvulsive Therapy

For severely depressed pregnant women electroconvulsive therapy (ECT) has been advocated by several researchers as an effective treatment option (Bhatia, Baldwin, & Bhatia, 1999; Dorn, 1985; Livingston, Johnstone, & Hadi, 1994; Rabheru, 2001; Walker & Swartz, 1994; Yellowlees & Page, 1990). There are no randomized trials.
## Table 2-6. Postpartum Depression Treatment Studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Design</th>
<th>Participants</th>
<th>Intervention</th>
<th>Outcome Measure</th>
<th>Results</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Appleby et al., 1997)</td>
<td>RCT(^1) Random allocation by computer-generated numbers Double-blinding Intent-to-treat</td>
<td>87 UK women with major PPD(^2) at 6 to 8 weeks postpartum</td>
<td>Four treatment groups: (1) fluoxetine and 1 CBT session, (2) fluoxetine and 6 CBT sessions, (3) placebo and 1 CBT session, or (4) placebo and 6 CBT sessions -Sessions derived for health visitors after brief training but provided by a psychologist with no previous clinical experience over 12 weeks</td>
<td>PPD at 1, 4, and 12 weeks post-treatment Clinical interview, EPDS, and HRSD</td>
<td>Significant improvements seen in all 4 treatment groups. The improvement with fluoxetine was significantly greater than placebo. Improvement after 6 sessions of counselling was significantly greater than after one session. Interaction between counselling and fluoxetine was not statistically significant. All group improvements were complete by 4 weeks.</td>
<td>Significant number of eligible women declined participation due to reluctance to take antidepressant medication No true control group (no treatment)</td>
</tr>
<tr>
<td>(Stowe et al., 1995)</td>
<td>Open-label single group</td>
<td>26 US women with major depression that developed within 24 weeks postpartum</td>
<td>8 weeks of sertraline using an initial dose of 50mg/day, adjusted according to side effects and depression severity, to a maximum dose of 200mg/day.</td>
<td>PPD post-treatment SIGH-D, EPDS, and BDI</td>
<td>20 out of 24 women exhibited a salutary response as defined by (&gt;50%) reduction in SIGH-D baseline scores; 14 out of 21 women demonstrated complete symptom remission.</td>
<td>Small sample size Lack of a control group Participants were not blinded to treatment Potential co-intervention through the provision of support</td>
</tr>
<tr>
<td>(Suri et al., 2001)</td>
<td>Open-label single group</td>
<td>6 US women with major PPD onset within the first 8 weeks - Identified using the EPDS and HRSD</td>
<td>Fluvoxamine treatment, 50 mg/day titrated to 150 mg/day over 2 weeks</td>
<td>PPD at 8 week post-treatment HRSD</td>
<td>Significant decline in HRSD scores over time with the greatest degree of improvement occurring between weeks 2 and 3</td>
<td>Small sample size Lack of a control group Participants were not blinded to treatment</td>
</tr>
<tr>
<td>(Cohen et al., 2001)</td>
<td>Open-label single group</td>
<td>15 US women who met DSM-III-R criteria for major depressive disorder with onset within the first 12 weeks postpartum</td>
<td>8 weeks of venlafaxine (immediate release; (M) dose = 162.5 mg/day)</td>
<td>PPD post-treatment HRDS</td>
<td>Twelve of 15 women experienced remission of major depression (HRSD score below 8)</td>
<td>Small sample size Lack of a control group Participants were not blinded to treatment</td>
</tr>
<tr>
<td>Study</td>
<td>Design</td>
<td>Participants</td>
<td>Intervention</td>
<td>Outcome Measure</td>
<td>Results</td>
<td>Limitations</td>
</tr>
<tr>
<td>---------------</td>
<td>----------------------</td>
<td>---------------------------------------------------</td>
<td>------------------------------------------------------------------------------</td>
<td>-----------------------------</td>
<td>------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------</td>
</tr>
</tbody>
</table>
| (Spinelli, 1997) | Single group pilot study | 13 US pregnant women Identified using a structured clinical interview for DMS-III-R and HRSD | 16-week treatment program incorporating weekly, 50-minute individual IPT sessions | PPD post-treatment and at 12 weeks postpartum HRSD, BDI, EPDS | Mean depression scores decreased significantly for all measures from week 0 to 16. Of the 10 women available at 12 weeks, none reported depressive symptomatology | Small sample size  
Lack of a control group  
Atypical sample (12 participants had a personal or family history of depression)  
Intervention provider was not reported |
| (Stuart & O'Hara, 1995b) | Single group | 6 US women (on average 4 months postpartum) Identified using DSM-III-R criteria | 12 weeks of IPT with modifications to include assistance with marital disputes and major role transitions | PPD post-treatment HRSD, BDI, and EPDS, | Significant changes for all measures were found post-treatment | Small sample size  
Lack of a control group  
Intervention provider was not reported |
| (O'Hara et al., 2000) | RCT Group allocation based on a random numbers table Power analysis Intent-to-treat | 120 US women  
Multi-stage community screening  
Identified using the IDD, SCID, and HRSD | Twelve 60-minute individual sessions by trained therapists during a 12-week period in standard fashion according to manual guidelines | PPD at 4, 8, 12 weeks following group assignment HRSD by interview BDI self-report | Recovery rates based on HRSD scores (< 7) significantly favoured IPT (37.5%) over the waiting list controls (13.7%). Based on BDI scores (<10), again recovery favoured IPT (43.8% vs. 13.7%) | Participants were mostly educated, Caucasian, married women  
Clinical interviewers were not blinded to group allocation |
| (Klier et al., 2001) | Single group | 17 Austrian women between 4 to 45 weeks postpartum presenting to a maternal mental health service either through referral or advertisement Identified using the DSM-IV and HRSD | Two 60-minute individual sessions to explain IPT, nine weekly 90-minute group sessions, one 60 minute individual termination session, and telephone numbers of other group members for support | PPD post-treatment and 6-month follow-up EPDS and HRSD | At post-treatment 10 of 17 women (59%) demonstrated full remission (HDRS < 9), 5 women (29%) demonstrated partial remission (score decrease >50%) and 2 women (13%) demonstrated no improvement. Women’s depression levels from post-treatment to 6-month follow-up were stable or continued to decrease. | Small sample size  
Lack of a control group  
Possible co-intervention through the provision of peer support outside of the group setting  
35% of participants terminated intervention early  
Investigator-based assessments of treatment outcome |
<table>
<thead>
<tr>
<th>Study</th>
<th>Design</th>
<th>Participants</th>
<th>Intervention</th>
<th>Outcome Measure</th>
<th>Results</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Appleby et al., 1997)</td>
<td>See above</td>
<td>37 Australian women screened by nurses identified using EPDS and clinical interview. 1 = 17 mothers  C = 20 mothers</td>
<td>Six weekly 60 minute home-based CBT sessions by trained Early Childhood Nurses (ECN)</td>
<td>PPD post-treatment and 6-month follow-up EPDS and MADRS</td>
<td>No significant group difference in mean EPDS or MADRS scores at all time periods</td>
<td>Small sample size  Inexplicit randomization process  Significant group differences in baseline EPDS 70% of control ECNs used some form of problem-solving and pleasant-event scheduling providing significant similarities to the intervention</td>
</tr>
<tr>
<td>(Prendergast &amp; Austin, 2001)</td>
<td>RCT</td>
<td>48 French women screened using EPDS verified using HRSD and BDI. 1 = 18 mothers  C = 30 mothers</td>
<td>Five to eight 1-hour weekly home-visits (M = 6.6, SD = 1.6) that had four components (supportive, educational, cognitive-behavioural, and psychodynamic)</td>
<td>PPD at post-treatment EPDS, HRSD, &amp; BDI</td>
<td>Significant group differences. Recovery rates (HRSD score below 7) were 66.6% for the intervention group versus 6.6% for the control group.</td>
<td>Small sample size  Non-random group allocation  High initial dropout after group assignment</td>
</tr>
<tr>
<td>(Chabrol et al., 2002)</td>
<td>Quasi-experimental Group allocation based on alternate numbers</td>
<td>20 Australian women with severe and long-standing PPD which developed within 6 months were recruited by local hospitals and maternal health centres. EPDS and BDI. 1 = 10 mothers  C = 10 mothers</td>
<td>Ten weekly 1.5 hour group sessions based on CBT conducted by a clinical psychologist. Women exchanged telephone numbers and met outside the program</td>
<td>PPD post-treatment EPDS, BDI and POMS</td>
<td>Following treatment, significant reductions in depression scores on all measures within the treatment group and between the treatment and control groups. However, due to initial severity of PPD, many women were still moderately depressed following treatment. Depression scores did not change over the 10 weeks for control group women.</td>
<td>Small sample size  Inexplicit randomization process  40% of women were on anti-depressant medication  40% of participants terminated intervention early  Possible co-intervention through the provision of peer support outside of the group setting</td>
</tr>
</tbody>
</table>

**Cognitive Behavioural Therapy**
<table>
<thead>
<tr>
<th>Study</th>
<th>Design</th>
<th>Participants</th>
<th>Intervention</th>
<th>Outcome Measure</th>
<th>Results</th>
<th>Limitations</th>
</tr>
</thead>
</table>
| (Fleming et al., 1992) | Quasi-       | 142 Canadian women recruited on postpartum wards to return screening instrument at 2 weeks Identified using CES-D and EPDS  
I₁ = 44 mothers  
I₂ = 15 mothers  
C = 83 mothers | Two treatment groups: (1) eight weekly semi-structured group sessions lasting 2 hours provided by 2 psychologists, (2) ‘group by mail’ where transcripts of the preceding support group were mailed to women | PPD at 6 weeks and 5 months CES-D | At 5 months, there was a significant improvement in maternal mood independent of group allocation. The supportive interventions did not modify maternal mood. Depressed mother had more negative feelings towards themselves, their partners, and the motherhood role than non-depressed women. | Non-random group allocation  
Significant differences in group sizes  
‘Depressed’ and ‘non-depressed’ women participated in all study groups  
Weak measure of PPD |
| (Chen et al., 2000)   | RCT          | 60 Chinese women recruited on postpartum wards to return screening instrument at 3 weeks Identified using BDI  
I = 30 mothers  
C = 30 mothers | Four weekly semi-structured group sessions lasting 1.5 to 2 hours, facilitated by a nurse | PPD post-treatment at 4 weeks BDI | Significant decrease in BDI scores in women attending support group. At the 4-week assessment 60% of women in the control group remained depressed in comparison to only 33% in the support group. | Only 44% of mothers returned screening questionnaire  
Inexplicit randomization process  
Unstandardized intervention  
Data analysis was not intent-to-treat |
| (Morgan et al., 1997) | Single group | 34 Australian women including 20 partners Identified using EPDS | Eight weekly 2-hour group sessions and 1 couple session facilitated by a nurse and occupational therapist using psychotherapy and cognitive behavioural strategies Telephone support from facilitators and referral available between groups if required | PPD post-treatment at 8 weeks and at 12 month follow-up EPDS and GHQ | Significant decrease in maternal scores pre and post-treatment. 22% of women scored > 12 on EPDS post-treatment and no women exhibited depressive symptoms at the 12 week follow-up. | Small sample size  
Atypical sample (74% had spent 1 week in a residential unit to help with mothering issues)  
Lack of a control group  
Co-interventions as 50% were receiving treatment by a health professional and ‘some’ were on anti-depressant medication |
<table>
<thead>
<tr>
<th>Study</th>
<th>Design</th>
<th>Participants</th>
<th>Intervention</th>
<th>Outcome Measure</th>
<th>Results</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Dennis, 2003)</td>
<td>Pilot RCT</td>
<td>42 Canadian women Screened by public health nurses during immunization clinic</td>
<td>Telephone-based support from a mother recruited from the community who previously experienced PPD and received a 4-hour training session Support individualized and based on maternal need</td>
<td>PPD at 4 and 8 weeks post randomization EPDS</td>
<td>Significant group differences were found in probable major PPD (EPDS &gt; 12) at all time periods. At the 4-week assessment 40.9% of women in the control group scored &gt; 12 on the EPDS in comparison to only 10% in the peer support group. Similar findings were found at 8 weeks.</td>
<td>Small sample size</td>
</tr>
<tr>
<td>(Misri, Kostaras, Fox et al., 2000)</td>
<td>RCT</td>
<td>29 Canadian women who met the DSM-IV criteria for major depressive disorder with postpartum onset</td>
<td>7 psychoeducational visits with a psychiatrist during which the mother’s partner participated in 4 of the 7 sessions</td>
<td>PPD post-treatment and 4 week follow-up EPDS</td>
<td>Immediately post-intervention there were no significant group differences in mean EPDS scores (p = 0.20). At the 4-week follow-up, significant group differences in mean EPDS scores favouring the intervention group were found (M = 8.6, SD = 5.2 vs. M = 14.7, SD = 7.2, p = 0.013).</td>
<td>Small sample size</td>
</tr>
<tr>
<td>(Holden et al., 1989)</td>
<td>RCT</td>
<td>50 UK women Community based EPDS screening at 6 weeks with a second screening at 13 weeks via psychiatric interview</td>
<td>8 weekly counselling visits at home by health visitors trained in non-directive counselling</td>
<td>PPD at 13 weeks post randomization EPDS and clinical interview</td>
<td>Significant group differences were found. According to RDC criteria, 18 (69%) of the 26 depressed women in the counselled group had fully recovered in comparison to only nine (38%) of the 24 women in the control group.</td>
<td>Small sample size</td>
</tr>
</tbody>
</table>

**Partner Support**

**Non-Directive Counselling**
<table>
<thead>
<tr>
<th>Study</th>
<th>Design</th>
<th>Participants</th>
<th>Intervention</th>
<th>Outcome Measure</th>
<th>Results</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Wickberg &amp; Hwang, 1996a)</td>
<td>RCT</td>
<td>31 Swedish women 2-stage population-based screening at 8 and 12 weeks using EPDS 1 = 15 mothers C = 16 mothers</td>
<td>6 weekly 1-hour counselling visits at home by nurses trained in non-directive counselling</td>
<td>PPD at 6 weeks post randomization Modified MADRS</td>
<td>Significant group differences were found. Twelve (80%) of 15 women with major depression in the study group were fully recovered after the intervention compared to four (25%) of 16 in the control group.</td>
<td>Small sample size Inexplicit randomization process</td>
</tr>
<tr>
<td>(Gregoire et al., 1996)</td>
<td>RCT</td>
<td>61 UK women with major depression, which began within 12 weeks postpartum and persisted for up to 18 months -Identified using EPDS and clinical interview 1 = 34 mothers C = 27 mothers</td>
<td>12 weeks of transdermal 17 beta-oestradiol 200 micrograms daily alone, then 12 weeks with added cyclical dydrogesterone 10mg daily for 12 days each month</td>
<td>PPD every 4 weeks for 24 weeks (end of treatment) EPDS and clinical interview</td>
<td>During the first 4 weeks of therapy, women receiving oestrogen improved significantly more (M = 13.3, SD = 5.7) than women in the placebo group (M = 16.5, SD = 5.3).</td>
<td>Small sample size 47% of women in the intervention group and 37% in the control group were taking antidepressant medication at trial enrolment A high EPDS cut-off score of 14 was used to determine initial eligibility Inexplicit randomization process</td>
</tr>
<tr>
<td>(Ahokas et al., 2001)</td>
<td>Open-label single group</td>
<td>23 Finnish women fulfilling ICD-10 criteria for major depression with postpartum onset</td>
<td>8-weeks of sublingual 17beta-estradiol</td>
<td>PPD at 2 weeks of treatment MADRS</td>
<td>MADRS scores were compatible with clinical recovery in 19 (82.6%) women.</td>
<td>Small sample size Lack of a control group</td>
</tr>
<tr>
<td>Study</td>
<td>Design</td>
<td>Participants</td>
<td>Intervention</td>
<td>Outcome Measure</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>--------</td>
<td>--------------</td>
<td>--------------</td>
<td>-----------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Relaxation/Massage Therapy</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Field et al., 1996)</td>
<td>RCT</td>
<td>32 depressed US adolescent women Identified using BDI Massage = 16 mothers Relaxation therapy = 16 mothers</td>
<td>30 minute massage per day on two consecutive days per week for five consecutive weeks 30 minute relaxation sessions per day on two consecutive days per week for five consecutive weeks</td>
<td>PPD at session 1 and 10 POMS 14 item depression subscale</td>
<td>Relaxation therapy had no effect on post-therapy depression scores at either session 1 or 10. However, massage therapy had a significant immediate effect on depression scores at both time periods. Small sample size Lack of a true control group Inexplicit randomization and trial procedures Weak measure of PPD Expressed disappointment may have influenced physiological results</td>
<td></td>
</tr>
<tr>
<td>(Onozawa et al., 2001)</td>
<td>RCT</td>
<td>34 UK primiparous women identified as being depressed at 4 weeks postpartum Identified using EPDS I = 19 mothers C = 15 mothers</td>
<td>Five weekly 1-hour infant massage classes and a 30-minute informal support group</td>
<td>PPD at last session EPDS</td>
<td>Median EPDS score for the massage group at the final session was 5.0 (95% CI = 8.0-14.2) in comparison to 10.0 (95% CI = 4.6 – 9.0) for the control group Small sample size Inexplicit randomization process High attrition rate with the massage group Lack of intent-to-treat</td>
<td></td>
</tr>
<tr>
<td><strong>Bright Light Therapy</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Corral, Kuan, &amp; Kostaras, 2000)</td>
<td>Case report</td>
<td>2 Canadian women with severe postpartum depression Identified using HRSD</td>
<td>Daily phototherapy by means of a 10,000-lux light box for 30 minutes for 4 weeks</td>
<td>PPD at last session at 4 weeks HRSD (29 item version)</td>
<td>HRSD scores dropped from above 29 and 28 to 11 and 12 respectively. Small sample size Lack of a control group</td>
<td></td>
</tr>
<tr>
<td>(Oren et al., 2002)</td>
<td>Single group</td>
<td>16 pregnant US women with diagnosis of major depression</td>
<td>Ultraviolet-screened diffuse white fluorescent light source incorporating a 100,000-lux box, tilted downward at home for 60 minutes daily beginning within 10 minutes of awakening for at least 3 weeks</td>
<td>Depression after 3 weeks of treatment HRSD</td>
<td>HRSD depression rating improved moderately by 49% after 3 weeks (between weeks 0 and 3 t = 6.27, p &lt; 0.001) Small sample size Lack of a control group</td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Design</td>
<td>Participants</td>
<td>Intervention</td>
<td>Outcome Measure</td>
<td>Results</td>
<td>Limitations</td>
</tr>
<tr>
<td>-----------------------</td>
<td>-------------------------</td>
<td>------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>-----------------</td>
<td>-------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>(Parry et al., 2000)</td>
<td>Single group</td>
<td>9 US women with DSM-IV criteria of depression with onset either in pregnancy or within 1 year postpartum</td>
<td>Either early-night sleep deprivation (ESD), sleep deprived in the early part of one night and slept from 03:00-07:00 h, or late-night sleep deprivation (LSD), deprived of sleep in the latter part of one night and slept from 21:00-01:00 h</td>
<td>PPD before and after the night of sleep deprivation and after a night of recovery sleep HRSD and BDI</td>
<td>More significantly participants responded to LSD (nine of 11 trials: 82%) compared with ESD (two of six trials: 33%) and they responded more after a night of recovery sleep (nine of 11 nights: 82%) than after a night of sleep deprivation (six of 11 nights: 55%).</td>
<td>Small sample size Lack of a control group Certain items were deleted because the results could be meaningfully related to the mother’s condition during a brief sleep deprivation protocol Limited number of women complied with the request to complete daily mood ratings</td>
</tr>
<tr>
<td>(Hiscock &amp; Wake, 2002)</td>
<td>RCT</td>
<td>156 Australian women of infants aged 6 to 12 months with severe sleep problems Subgroup of women categorised as depressed Identified using EPDS I = 78 mothers C = 78 mothers</td>
<td>Discussion on behavioural infant sleep intervention (controlled crying) delivered over three consultations with a pediatric trainee</td>
<td>PPD at 8 and 16 weeks post-randomization EPDS</td>
<td>For the subgroup of women with baseline EPDS scores of above 9, depression scores fell significantly further for women in the intervention group than the control group at 8 weeks (-6.0 vs. -3.7, p = 0.01) and at 16 weeks (-6.5 vs. -4.2, p = 0.04).</td>
<td>67% of eligible mothers accepted trial participation 56% of participants had EPDS scores &lt; 10 at trial the start of the trial making a significant reduction in scores unlikely</td>
</tr>
</tbody>
</table>

1 RCT = randomized controlled trial; 2 PPD = postpartum depression; 3 I = intervention group and C = control group
Implications for Policy, Practice, and Research

While limited research has been conducted on the efficacy of pharmacological interventions for the specific treatment of postpartum depression, four small studies have shown antidepressant medications, especially SSRIs, may have a therapeutic effective for severely depressed women. Many women find this an unattractive treatment option (Appleby et al., 1997; Whitton et al., 1996). As such, high attrition of participants or lack of intervention compliance may be of concern in future studies requiring randomization to treatment conditions. Building upon the primarily descriptive studies thus far, well-designed randomized controlled trials are needed to significantly advance this postpartum depression treatment approach. However, it is important to highlight that antidepressant medication has been shown to be highly effective in the treatment of general depression. As such, a recommendation for its use to treat postpartum depression could be made based on this empirical work.

Research into the characteristics of pregnant and postpartum women and their intervention choice (e.g., antidepressant medication vs. other options) would be beneficial to health professionals. Furthermore, while no specific pharmacological treatment guidelines are available for postpartum depression, primarily due to the lack of research, there is a tendency to treat postpartum depression with less intensity (i.e., lower dose of medication and duration of treatment) than general depression (Jermain, 1995; Nonacs & Cohen, 1998, 2002). As such, rapidity of response to different antidepressant medications requires further investigation. Even with the potential benefits of antidepressant medication, it is important to note that this biological treatment approach primarily addresses symptom reduction and may not assist in altering the psychological, family and social factors that often contribute to or maintain the depression.

There are several different psychological approaches to the treatment of postpartum depression including cognitive behavioural therapy and interpersonal psychotherapy. Cognitive behavioural therapy is also an effective treatment for general depression with a meta-analysis of 28 studies suggesting that, given over an average of 14.9 weeks, this intervention is as effective (Dobson, 1989) as medication or other psychotherapies. In a large recent analysis of 4 trials (DeRubeis, Gelfand, Tang, & Simons, 1999), cognitive behavioural therapy fared as well as antidepressant medication with ‘severely’ depressed outpatients in four major comparisons. However, considerable time, commitment and cost is required from cognitive behaviour therapy participants and approximately 10% to 40% fail to complete full treatment, a compliance rate similar to pharmacotherapy (Evans et al., 1992). In this current review, four trials were found evaluating the effectiveness of cognitive behavioural therapy related to postpartum depression but all suffered significant methodological limitations, such as a small sample size or lack of a true control group. At this time, there is poor evidence regarding the inclusion or exclusion of this approach in postpartum depression treatment programs, but the primarily beneficial results suggest that further research is warranted. Similarly, psychotherapies that target interpersonal and/or current psychological problems related to depression have
been shown to be more effective than long-term analytic psychotherapies (Elkin et al., 1989). In this review, four studies were found evaluating the effectiveness of interpersonal psychotherapy; however, only one investigation was a well-designed trial (O'Hara et al., 2000). The results from this trial and the other single group studies suggest there is some evidence to support the recommendation that this approach may be effective in the treatment of postpartum depression. As such, structured cognitive behavioural therapy and interpersonal psychotherapy holds promise and well-designed trials with large samples are warranted. Future investigations should include long-term follow-up after intervention discontinuation and be designed to determine the comparative effectiveness of pharmacological and psychological treatments, using trained health professionals and standardized interventions.

Research has clearly shown that a lack of social support is a significant predictor of postpartum depression. As such, peer support interventions potentially have beneficial effects in treating women who have mild to moderate depression. Three studies have been found evaluating the effectiveness of professionally facilitated support groups. Unfortunately, theoretical limitations, such as the inclusion of both depressed and non-depressed women, and methodological weaknesses, including small samples sizes and single group or non-random samples, render the results equivocal. Well-designed trials with large, homogeneous samples are warranted. Future research should also include self-help groups (i.e., groups not facilitated by a health professional) to extend the testing of lay support models with mild to moderately depressed women and evaluations of eligible mothers who decline group interventions should be conducted to identify help-seeking barriers and possible solutions. Evaluations of the group intervention should include measures that assess group dynamics, social comparisons, and the provision of peer (mother-to-mother) support to determine the salutary components of support groups. A new intervention that holds promise is telephone-based peer support and further research is warranted.

One area that has received little attention is the role the spouse or partner plays in the prevention of or recovery from postpartum depression. Partners can be an excellent source of instrumental (e.g., sharing of childcare and domestic responsibilities) and emotional support and can be a mediating link between the mother and family members who may not understand the nature of postpartum depression. Further research is needed to identify the type and amount of social support that is most beneficial is assisting with postpartum depression.

Two European trials have been conducted evaluating the effectiveness of non-directive counselling with positive results suggesting this treatment modality may be a viable option for mothers with mild to moderate postpartum depression. These trials have demonstrated the feasibility of population-based screening and the application of home visiting using trained health professionals. Unfortunately, the most immediate problem is the small sample size in both trials. Contextual factors also decrease the application of the results to a
North American population where differences in the delivery of postpartum care exist. As such, a large randomized controlled trial is needed to replicate these auspicious results.

Oestrogen therapy has been advocated with preliminary results from two studies demonstrating effectiveness. Until better controlled trials are conducted, it is unclear whether specific subgroups of mothers, especially those with treatment resistant depression, derive an antidepressant benefit from supplemental oestrogen. Further research is highly recommended to establish dose response relations, optimum treatment duration, as well as the antidepressant mechanisms. Whether breastfeeding women can use oestrogen must also be carefully examined and research is needed into how changes in sex-steroid concentrations contribute to the occurrence of postpartum depression (Gregoire et al., 1996).

Two small studies have evaluated the effect of massage therapy on maternal mood demonstrating positive results. However, severe methodological limitations in both studies render the effect unknown. Maternal/infant massage therapy and sleep interventions hold promise and, with further research, these interventions may be beneficial secondary treatment options. Finally, for severely depressed individuals with acute suicidality or psychosis, electroconvulsive therapy is frequently the treatment of choice (Nonacs & Cohen, 2002). However, the relative effectiveness of electroconvulsive therapy for severely depressed expectant or postpartum women is unknown as no randomized controlled trials exist for this indication and most of the research thus far has been case studies. Similarly, only two case studies have been found exploring the effect of bright light therapy on severely depressed mothers. While these treatment approaches are not first-line options, if they are to become a component of a multifactorial treatment program well-designed randomized controlled trials are required to ascertain whether maternal mood improvement is specific to the intervention or a placebo effect from open treatment.

This review has clearly demonstrated that postpartum depression presents many special methodological complexities that need to be considered if scientific knowledge is to progress. First, there are particular difficulties in defining the target group to be studied, as diagnosis is much less concrete than in other areas where an initial assessment can be confirmed by laboratory tests. Second, many of the treatments used are hard to define with clarity as psychological and psychosocial interventions often involve talking and manipulation of the environment. Replicating such treatment with fidelity is challenging. Third, the nature of the interventions employed frequently result in co-interventions. Fourth, there are difficulties in establishing the relative costs and benefits of treatment, arising from the relapsing/remitting nature of postpartum depression. Finally, the context of postpartum depression research is crucial and the social, cultural, and organizational environment in which postpartum depression services takes place is highly variable. For example, the same intervention can have differing effects depending on context and variations in the control group.
Many of the dilemmas with postpartum depression research begin from the way in which interventions are evaluated. In this review, it was found that there was limited agreement on outcome measures, although the EPDS was the most consistently used measure of depressive symptomatology. Most studies obtained no information on maternal perceptions, such as whether the women simply felt better or even liked the intervention. Although postpartum depression can occur within the first year, most trials had follow-up periods of less than 6 months. Of the trials conducted, most were small with a mean sample size of approximately 43 women, although 300 would be more appropriate to detect clinically significant changes in depressive symptomatology. There were also high attrition rates, especially with group interventions. Examination of the wider impact of postpartum depression through economic evaluations was rarely conducted and impossible to complete post hoc due to small sample sizes. In addition, the practice of excluding participants that might potentially benefit (e.g. family history of depression) reduces generalizability. Finally, little attention has been paid to the context in which postpartum depression interventions have been evaluated. This covers not only the broad social and policy context of different countries but also control groups, which can be more variable than the intervention studies.

The challenge is to conduct methodologically rigorous randomized controlled trials remembering that one expensive randomized controlled trial may prove more cost effective than a large number of small studies with no meaningful results. To ensure that trials are well designed the following points need to be considered. Difficulties in definition of the postpartum depression should be confronted by using structured diagnoses or psychometrically tested self-report instruments such as the Edinburgh Postnatal Depression Scale. Dialogue between researchers should be encouraged to promote a consistency in outcome measures and research methods. Adequate sample sizes based on power analyses should be incorporated such that the results can be compared across different postpartum samples. Researchers should consider multiple dimensions of improvement. However, trials should focus on a small number of clear outcomes, in the interest of both clarity and maintaining the involvement of women. Long-term effects should be addressed by adequate length of follow-up. Trials should also be analysed by intention-to-treat without excluding those who dropout due to a change in treatment. Intervention replication can be achieved through a concise account not only of the intended but also the actual intervention in both the experimental and control group. Finally, maternal evaluations should be included to understand the nature of the intervention as well as what are important outcomes.

At present, definite conclusions cannot be reached about the relative effectiveness of these different treatment approaches due to the lack of well-designed investigations. Randomized controlled trials with large and representative samples are needed to compare different treatment modalities, examine the effectiveness of individual treatment components, and determine which treatments are most useful for women with different risk factors or clinical presentations of postpartum depression. As there is no single etiological
pathway by which women develop postpartum depression, it is improbable that a single treatment modality will be effective for all women. A multifactorial treatment approach, which combines the contributions of the psychological, psychosocial, and biological factors, is likely to be most beneficial as it recognizes various etiological factors and individual variations.

Table 2-7. Summary Recommendations for Treatment Interventions

<table>
<thead>
<tr>
<th>Intervention Strategy</th>
<th>Study</th>
<th>Research Design Rating</th>
<th>Quality Rating</th>
<th>Classification of Recommendation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pharmacological</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Antidepressant Medication</td>
<td>(Appleby et al., 1997)</td>
<td>RCT: I</td>
<td>Fair</td>
<td>I4</td>
</tr>
<tr>
<td></td>
<td>(Stowe et al., 1995)</td>
<td>Descriptive: III</td>
<td>Poor</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(Suri et al., 2001)</td>
<td>Descriptive: III</td>
<td>Poor</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(Cohen et al., 2001)</td>
<td>Descriptive: III</td>
<td>Poor</td>
<td></td>
</tr>
<tr>
<td><strong>Psychological</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interpersonal Psychotherapy</td>
<td>(Spinelli, 1997)</td>
<td>Descriptive: III</td>
<td>Poor</td>
<td>I</td>
</tr>
<tr>
<td></td>
<td>(Stuart &amp; O'Hara, 1995b)</td>
<td>Descriptive: III</td>
<td>Poor</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(O'Hara et al., 2000)</td>
<td>RCT: I</td>
<td>Fair</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(Klier et al., 2001)</td>
<td>Descriptive: III</td>
<td>Poor</td>
<td></td>
</tr>
<tr>
<td>Cognitive Behavioural Therapy</td>
<td>(Appleby et al., 1997)</td>
<td>RCT: I</td>
<td>Fair</td>
<td>I4</td>
</tr>
<tr>
<td></td>
<td>(Prendergast &amp; Austin, 2001)</td>
<td>RCT: I</td>
<td>Poor</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(Chabrol et al., 2002)</td>
<td>Quasi-Experimental: II-1</td>
<td>Poor</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(Meager &amp; Milgrom, 1996)</td>
<td>Pilot RCT: I</td>
<td>Poor</td>
<td></td>
</tr>
<tr>
<td><strong>Psychosocial</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peer Support</td>
<td>(Fleming et al., 1992)</td>
<td>Quasi-Experimental: II-1</td>
<td>Poor</td>
<td>I</td>
</tr>
<tr>
<td></td>
<td>(Chen et al., 2000)</td>
<td>RCT: I</td>
<td>Poor</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(Morgan et al., 1997)</td>
<td>Descriptive: III</td>
<td>Poor</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(Dennis, 2003)</td>
<td>Pilot RCT: I</td>
<td>Fair</td>
<td></td>
</tr>
<tr>
<td>Partner Support</td>
<td>(Misri et al., 2000)</td>
<td>RCT: I</td>
<td>Poor</td>
<td>I</td>
</tr>
<tr>
<td><strong>Non-Directive Counselling</strong></td>
<td>(Holden et al., 1989)</td>
<td>RCT: I</td>
<td>Fair</td>
<td>I</td>
</tr>
<tr>
<td></td>
<td>(Wickberg &amp; Hwang, 1996a)</td>
<td>RCT: I</td>
<td>Fair</td>
<td></td>
</tr>
<tr>
<td><strong>Hormonal</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oestrogen Therapy</td>
<td>(Griepio et al., 1996)</td>
<td>RCT: I</td>
<td>Poor</td>
<td>I</td>
</tr>
<tr>
<td></td>
<td>(Ahokas et al., 2001)</td>
<td>Descriptive: III</td>
<td>Poor</td>
<td></td>
</tr>
<tr>
<td>Intervention Strategy</td>
<td>Study</td>
<td>Research Design Rating&lt;sup&gt;1&lt;/sup&gt;</td>
<td>Quality Rating&lt;sup&gt;2&lt;/sup&gt;</td>
<td>Classification of Recommendation&lt;sup&gt;3&lt;/sup&gt;</td>
</tr>
<tr>
<td>-----------------------</td>
<td>-------</td>
<td>-------------------------------</td>
<td>-----------------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>Relaxation/ Massage Therapy</td>
<td>(Field et al., 1996)</td>
<td>RCT: I</td>
<td>Poor</td>
<td>I</td>
</tr>
<tr>
<td></td>
<td>(Onozawa et al., 2001)</td>
<td>RCT: I</td>
<td>Poor</td>
<td></td>
</tr>
<tr>
<td>Bright Light Therapy</td>
<td>(Corral, Kuan, &amp; Kostaras, 2000)</td>
<td>Case Report: III</td>
<td>Poor</td>
<td>I</td>
</tr>
<tr>
<td></td>
<td>(Oren et al., 2002)</td>
<td>Case Report: III</td>
<td>Poor</td>
<td></td>
</tr>
<tr>
<td>Sleep Interventions</td>
<td>(Parry et al., 2000)</td>
<td>Descriptive: III</td>
<td>Poor</td>
<td>I</td>
</tr>
<tr>
<td></td>
<td>(Hiscock &amp; Wake, 2002)</td>
<td>RCT: I</td>
<td>Fair</td>
<td></td>
</tr>
</tbody>
</table>

1 I = evidence from randomized controlled trial(s); II-1 = evidence from controlled trial(s) without randomization; II-2 = evidence from cohort or case-control analytic studies, preferably from more than one centre or research group; II-3 = evidence from comparisons between times or places with or without the intervention, dramatic results in uncontrolled experiments could be included here; III = opinion of respected authorities, based on clinical experience, descriptive studies or reports of expert committees.

2 Good = a study (including meta-analyses or systematic reviews) that meets all design-specific criteria well; Fair = a study (including meta-analyses or systematic reviews) that does not meet (or it is not clear that it meets) at least one design-specific criterion but has no known “fatal flaw”; Poor = a study (including meta-analyses or systematic reviews) that has at least one design-specific “fatal flaw”, or an accumulation of lesser flaws to the extent that the results of the study are not deemed able to inform recommendation.

3 A = there is good evidence to recommend this approach; B = there is fair evidence to recommend this approach; C = the existing evidence is conflicting and does not allow making a recommendation for or against use of this approach, however other factors may influence decision-making; D = there is fair evidence to recommend against this approach; E = there is good evidence to recommend against this approach; I = there is insufficient evidence (in quantity and/or quality) to make a recommendation, however other factors may influence decision-making.

4 There is evidence based on the general depression research to recommend this approach.
References


