Childhood Maltreatment and Postnatal Depression: are there distinctive risks?

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor in Philosophy by

Brigitta Claudia Bende

March 2005

Contents

Tables

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Chapter 7

Appendix D

Figures

Acronyms

AND	Antenatal depression
PND	Postnatal depression
PrevD	Previous depression
PrevD:NIR	Previous depression – no impairment or referral
PrevD:IR	Previous depression – impaired or referred
NPrevD	No previous depression
CSA	Childhood Sexual Abuse
pp	Postpartum
SD	Standard deviation
OR	Odds' ratio
CI	Confidence Interval
PBI	Parental Bonding Instrument
BDI	Beck Depression Inventory
EPDS	Edinburgh Postnatal Depression Scale
PI	Predictive Index
SCID	Structured Clinical Interview for DSM
DSM	Diagnostic and Statistical Manual of Mental Disorders
ICD	International Classification of Diseases
ANC	Antenatal Clinic
CMW	Community Midwife
GP	General Practitioner
REC	Research Ethics Committee
R&D	Research and Development
LWH	Liverpool Women's Hospital
NHS	National Health Service

Acknowledgements

I am thankful to many people who over the years supported and sustained me in my endeavour.

I am particularly grateful to Professor Jonathan Hill for overall supervision and Dr Helen Sharp for her guidance and practical support. Feedback at PhD seminars at the Department of Psychiatry – Child Mental Health; presentations at the Department of Primary care and the School of Health Care Science have been very helpful.
Dr Philip Bell helped with interviews; Hazel Fothergill and Jeannette Chamberlain provided administrative assistance. Dr John Davies, University of Liverpool, spent many hours helping me to set up an Access Database, thereby ensuring a smooth administration of the study.

This study could not have been done without the tremendous support from a wide range of staff at the Liverpool Women's Hospital (LWH). In particular, I would like to thank all the Community Midwives, who recruited their patients into the study, often in difficult conditions. Sarah Drummond and Grace Edwards helped to obtain vital information. Support from senior management in the obstetric and midwifery services was crucial. Mr Steve Walkinshaw, Clinical Director of Obstetrics at Liverpool Women's Hospital, provided invaluable support and guidance throughout the study.

This study was conducted whilst I was in receipt of a training fellowship from the NHS Executive North West (R&D Fellowship number: RDO/33/77).
Financial assistance for the study administration has been provided by Liverpool Women's Hospital.

Finally, without Ron Price's tremendous editing of my Teutonic English, this work would be a far less enjoyable read.

Abstract

Brigitta Claudia Bende – Childhood Maltreatment and Postnatal Depression: are there distinctive risks?

Background:

There is convincing evidence that adverse childhood experiences such as parental low care and high control, childhood sexual abuse (CSA) and physical abuse contribute to the risk of depression in adult life. However, it is not known whether these risks operate in a similar manner in postnatal depression (PND). The study of childhood antecedents of PND can help to illuminate the mechanisms in PND and permit further exploration of similarities and differences between PND and depression at other times in women's lives.

Objectives:

This study set out to examine (1) whether childhood adversities (such as parental low care and high control; childhood sexual abuse (CSA) and physical abuse) are predictive of PND; (2) whether risks are distinctive and differ from both previous depression (PrevD) and antenatal depression (AND); (3) whether there is an independent association of childhood adversities with PND that is not mediated via a history of depression (PrevD) or AND; and (4) whether risks from adverse childhood experiences for newly developing PND differ from both those in the antenatal period and those occurring prior to pregnancy. Each of these questions were examined contrasting transient and persistent PND.

Methods:

A questionnaire-based study of a consecutive cohort of primiparous pregnant women was conducted with follow up at 6 and 12 weeks postpartum (pp).
At baseline, 1,029 women were recruited, of whom 821 women (80%) returned the antenatal, and 695 (68%) returned the six weeks postnatal measure. Of 733 women who had received the 12 weeks pp measure, 506 women returned this at time 3, representing a response rate of 74%.
Antenatal assessment consisted of antenatal depression (AND) using the Edinburgh Postnatal Depression Scale (EPDS); previous depression (PrevD) using the Predictive Index (PI); parental care and control using the Parental Bonding Instrument (PBI); and questions on childhood sexual and physical abuse used in other studies. Postnatal depression was assessed with the EPDS at 6 and 12 weeks pp.

Results:

Overall childhood risks of postnatal depression resembled those of depression in women at other times of their lives. The clear exception to this was CSA, which showed a lack of association. Childhood predictors were also associated with antenatal depression and previous depression in this sample of women. Childhood predictors of PND did not remain significant once the association with antenatal and previous depression with PND was accounted for. There were no indications that new-onset PND differed from PND generally. Childhood adversity, AND and PrevD were strongly associated with PND that persisted from 6 to 12 weeks pp. This was in marked contrast to weak associations with depression that had remitted by 12 weeks pp.

Conclusions:

The findings do not support the argument that PND is distinctive from depression in women in general. The most powerful predictor of PND, and particularly of PND that persists, is a history of depression and antenatal depression. The absence of an association between PND and CSA may point towards an activation of different risks at different times in women's lives.

Chapter 1

Introduction

This study is concerned with the relationship between antenatal (AND) and postnatal depression (PND) and childhood adversity. It is hoped that findings will inform further attempts at understanding transgenerational processes of perinatal mental ill health and its effects on mother–infant relationships and on infant development.

The last three decades have seen an increase in research into PND, despite a broadening acceptance that PND might not differ greatly from depression in women outside of the postpartum period, a trend underlined by the omission of the diagnosis in ICD10 (WHO 1992) and DSM-IV (American Psychiatric Association [hereafter APA] 1994).

What is it, then, that sustains the interest of part of the research community in the nature and the causes of PND? The answer is, in part, that there are indications of PND's distinctiveness – for example, the risk of recurrence of PND following post-natal depression appears to be specific in that once it has been experienced, there is a real risk that the sufferer is subject to a recurrence in subsequent pregnancies (Cooper & Murray 1995); while the possibility that PND poses a risk to early infant development is also a central concern (e.g. Cogill *et al.* 1986; Hay & Kumar 1995; Sharp *et al.* 1995; Murray *et al.* 1999; Cooper *et al.* 2003).

This consideration is of particular relevance to the study here reported. Findings of association between PND and impaired infant development are insufficient to establish a causal link; an alternative explanation is that factors which confer

vulnerability to depression in the mother also affect the infant. Out of several possibilities, the one considered here is that aspects of the mother's own experience of parenting may be of particular relevance to the likelihood that she will become depressed following the birth of her child and those same factors might also influence her capacity to care for her infant.

The notion that PND and depression in women outside of the puerperium are distinct has been disputed in a number of studies – in terms of prevalence (e.g. O'Hara *et al.* 1990; Cox *et al.* 1993), risks and symptom profiles (Whiffen & Gotlib 1993), though evidence that some women are at risk of recurrence specifically of PND in subsequent pregnancies (Cooper & Murray 1995) points to PND's distinctiveness. This is in line with findings of research into depression in women in general, which supports the idea that depression is a heterogeneous disorder in which a shared symptom profile is reached via distinct pathways.

A further reason for the focus on depression in the postnatal period is afforded by findings indicating that infants of women affected show adverse outcomes (Cogill *et al.* 1986; Hay & Kumar 1995; Sharp *et al.* 1995; Murray *et al.* 1999; Cooper *et al.* 2003), giving rise to a central concern about a possible critical period during which the quality of parenting and/or mother–infant interaction impacts significantly on the infant's long-term development at a time of particular vulnerability and dependence (Hay & Kumar 1995; Sharp *et al.* 1995; Campbell & Cohn 1997). In particular, the finding that infants of postnatally depressed mothers are at risk of being adversely affected raises the possibility that there are transgenerational issues at work which impact on future generations. These considerations underscore the importance of studying PND in its own right.

Despite numerous studies comparing PND and depression in women outside of the puerperium, some of the main findings within the general research into depression have rather surprisingly not been transferred to research of PND; specifically, early childhood adversities, which have now been clearly established as risks for depression in women (Bifulco *et al.* 1991; Mullen *et al.* 1993; Parker *et al.* 1995; Fergusson *et al.* 1996a) have not been investigated. This is of particular relevance to

the issue of the distinctiveness of postnatal depression: if postnatal depression has childhood antecedents different from depression occurring at other times, that would suggest distinct mechanisms. The various contributions of early adversity and other risks to the development of depression may be pointers to depression that requires to be conceived of as a heterogeneous rather than uniform disorder. Understanding PND with regard to processes involved in its causation may provide part of the answer to this heterogeneity jigsaw. For this, it would be particularly important to understand mechanisms of transgenerational processes, ranging from the mother's own childhood experiences to those of her pregnancy and the new baby.

In summary, research into theoretical models of and pathways to PND is important in understanding the specificity *and* the heterogeneity of depression. This study aims to clarify the relationship between AND and PND and childhood adversities. The section following provides an overview of relevant research findings on PND and antenatal depression to set the study in a context of established risk factors; chapter 2 reviews relevant research findings on depression in women across other times in the life-cycle, particularly concentrating on findings regarding the role of childhood adversity in relation to depression in general. Contributions on childhood adversity will then be discussed with reference to PND.

1.1 PND

PND is defined as a non-psychotic depressive episode encountered following the delivery of a child. Approaches to the definition of 'postnatal period' vary from one month – as required by DSM-IV (APA 1994) if specifying 'postnatal onset' – to the first postnatal year (e.g. Cooper *et al.* 1988). The most commonly used diagnostic systems such as DSM-IV (APA 1994) and ICD 10 (WHO 1992) do not contain specific diagnostic criteria for PND, though they include it within depressive disorder as a whole, with DSM-IV (APA1994) allowing for specification of postpartum onset within 4 weeks of delivery; ICD 10 (WHO 1992) allows only for classification in the section of puerperium-associated disorders if onset is within 6 weeks of delivery and the disorder does not meet criteria elsewhere classified.

A DSM-IV (APA 1994) diagnosis of major depression requires the presence of either of 2 main symptoms (depressed mood or anhedonia) and any of 4 of 7 other symptoms, which have to be present for at least 2 weeks for most of the day, most of the time, and affect functioning. Inclusion of postpartum depressive disorders within the general classification of depressive disorder means that assessments and criteria are not adjusted to a woman's postnatal status, which in itself may be directly responsible for changes resembling those of depressive symptoms. This may pose difficulties for detecting true depression, and hence these considerations have led to the development of specific screening tools for PND – such as the Edinburgh Postnatal Depression Scale (EPDS; see Cox *et al.* 1987) – which reduce the risk of postpartum status-confounding symptoms of depression.

Suggestions are due to be put forward for the inclusion of puerperium-specific disorders in the revised versions of DSM or ICD; extant proposals are to create a puerperal category with onset within the first three months of delivery, to include the range of sub-syndromal and mild to psychotic puerperal disorders (Cox 2004).

Assessment of PND involves either a clinical interview, concluding with a categorical diagnosis of PND, or an endorsement of symptoms on a self-report measure, such as the EPDS (Cox *et al.* 1987), which provides a measure of the severity of depressive symptoms rather than a diagnosis.

Before continuing with further detailed descriptions of postnatal and antenatal depression I briefly outline how they tend to be measured in research studies, thereby providing the rationale for the approach taken in this study. The section which follows therefore turns to questionnaire-based measures.

1.1.1 Measurement issues

1.1.1.1 Questionnaire-based measures

Early studies used a variety of measures to assess PND. Scales validated on non-puerperal samples were used by many, including the Beck Depression Inventory (BDI; see Beck *et al.* 1961), the Hamilton Depression Rating Scale (HRDS; see

Hamilton 1960), the Hospital Depression and Anxiety Scale (HDAS; see Zigmond *et al.* 1983), and posed a risk of misdiagnosing depression by including symptoms related to normal changes in the puerperium (such as insomnia) rather than to depression. Additionally, the variety of measures used compromised the ability to compare findings across studies. The development of the EPDS – specifically designed with puerperal women in mind and validated postnatally – addressed these difficulties, and this tool has become the most commonly used screening device in most countries.

Although paper and pencil self-report rating scales are generally simple to administer, particularly in large general population studies, and tend to yield good response rates (Murray & Carothers 1990; Cox *et al.* 1993), they are not without limitations, as they rely on the subject's understanding of the items. Equally, there has been concern regarding structured diagnostic interviews: Brugha *et al.* (1999) reported substantial differences in case findings between different diagnostic interviews. Thus there is no 'gold standard' for the diagnosis of mental disorders. Particularly for large community studies, a balance needs to be struck to achieve the requisite sensitivity and specificity for maximising the detection of depression in women without over-inclusively identifying as depressed women who in fact are not or who do not suffer from impairment. These considerations are outlined below with regard to the EPDS.

1.1.1.1.1 The EPDS

The EPDS, a ten-item self-administered questionnaire, was specifically developed by Cox *et al.* (1987) to screen for PND and to minimise the likelihood of including symptoms occurring in the puerperium that do not arise from depression. Cox *et al.* (1987) published validation data comparing the EPDS to Goldberg's Standardised Psychiatric Interview (SPI; see Goldberg *et al.* 1970) and Research Diagnostic Criteria (RDC; see Spitzer *et al.* 1978). Their original validation study using data from 84 mothers showed the EPDS, at a cut-off of 12/13 (out of a maximum of 30), to have a sensitivity (proportion of depressed women correctly identified, or true positives) of 86 per cent; specificity (proportion of non-depressed women correctly identified, or true negatives) of 78 per cent; and a positive predictive value

(proportion of those identified as depressed who *are* truly depressed) of 73 per cent for major and probable major depression.

Two further validation studies conducted in the UK differed in their conclusions regarding how well the EPDS detected depression, but generally agreed on its utility: the study by Harris *et al.* (1989a) of 147 women suggested a higher sensitivity with 95 per cent and a specificity of 93 per cent, which compared particularly well with Beck's Depression Inventory (BDI; see Beck *et al.* 1961) that had shown a sensitivity of only 68 per cent and a specificity of 88 per cent for major depression; the second validation study, carried out by Murray & Carothers (1990), involving a larger community sample of 702 postnatal women, found a lower sensitivity and a positive predictive value at the suggested cut-off of 12/13 of 67.7 and 66.7 per cent, respectively, with a specificity of 95.7 per cent for major and minor depression. However, there were no indications that altering the cut-off gave improved positive prediction rates.

With regard to AND, the original validation study by Murray & Cox (1990) on 100 antenatal women suggested a cut-off of 12/13 to detect major and minor depression at a sensitivity of 64 per cent, a specificity of 90 per cent, and a positive predictive value of 50 per cent. The use of a higher cut-off at 14/15 has been described as optimal for the detection of RDC (Spitzer *et al.* 1978) major depression, in that sensitivity was 100 per cent, a specificity of 96 per cent and a positive predictive value of 60 per cent (Murray & Cox 1990). There is general agreement that the EPDS tends to produce high response rates even with postal administration, with general acceptability to postnatal women.

The above considerations have led to the EPDS now being the most commonly used tool for screening both for PND and AND (Austin & Lumley 2003): most large studies have adopted this approach (e.g. Forman *et al.* 2000; Webster *et al.* 2000; Evans *et al.* 2001) and the 12/13 cut-off tends to be the most widely used.

For practical reasons, and because the EPDS has been shown in a meta-analysis to yield similar prevalences of PND as interview-based measures (O'Hara & Swain

1996), this study used the EPDS as a first-line measure, with SCID interviews as a validation of the diagnoses in a random sub-sample (see 3.5.1.1).

1.1.2 Prevalence of PND

1.1.2.1 Prevalence studies: methodological issues

Prevalence of PND varies according to the method employed. This can affect the comparison of findings from different studies and populations. Factors of importance here are:

- *The definition of PND*: the classification system used, the chosen cut-off of measures and the minimum duration of symptoms required for a diagnosis are likely to affect prevalence rates.
- *The period assessed*: distinct measures assess different time periods: whereas the EPDS enquires about symptoms occurring during the previous 7 days, the SCID assesses the previous 30.
- *The tools used to assess depression*: self-assessment questionnaires, used as screening tools in large general population studies, provide an *assessment of symptom severity* rather than a diagnosis, whereas interview-based assessments tend to provide a categorical diagnosis of the absence or presence of disorder, made by an observer. If tools are used that have not been developed specifically for the postnatal period, such as the BDI (Beck *et al.* 1961), prevalence rates of depression may not be reflective of the *true* prevalence of the disorder due to the inclusion of symptoms associated with normal puerperal physiological changes (Harris *et al.* 1989a).
- *The timing of the onset of depression*: depression might start in the antenatal and continue into the postnatal period; its onset also can antedate pregnancy with PND as a recurrence of prior depression. A narrower definition of 'new-onset PND' would specify an onset within the puerperium.
- *Defining the 'postnatal period' and the timing of assessments*: the extent of the postnatal period varies between studies, ranging from 4 weeks to 12 months. Moreover, assessments performed too early in the puerperium risk falsely including women suffering from maternal 'blues', a transient self-limiting mood disorder in the first week postpartum.

- *Methods of sampling*: Studies employing random or convenience sampling – e.g. clinical or referred groups – are likely to establish different prevalences and may affect any associations found between PND and factors of interest. Therefore, the findings of studies using convenience samples or of particular sub-groups of the general population may not be generalisable to the wider population.
- *Inclusion and exclusion criteria for the group under investigation*: many studies concentrate on primiparous women to reduce potential confounders, thus elevating the likelihood of findings being inapplicable to all women of childbearing age.

For a more exhaustive discussion of issues applying to retrospective methods in general, see section 2.2.1.

The section which follows summarises recent studies of the prevalence of PND, concentrating on meta-analytic and prospective approaches, which provide the most robust findings. Certain controlled studies, which also refer to prevalences, are described in later sections.

1.1.2.2 Prevalence studies: findings

A meta-analysis based on 59 studies involving 12,810 subjects by O'Hara & Swain (1996) reports prevalence rates of PND and risk factors using rigorous criteria for the inclusion of studies. The inclusion criteria relevant to prevalence were:

- employment of random or quasi-random sampling techniques;
- assessment of depression at least two weeks postpartum; and
- use of a validated or standardised measure.

This meta-analysis found an overall pooled prevalence of PND of 13 per cent (95 per cent; CI 12.3–13.4 per cent), with self-report measures providing a slightly higher prevalence of 14 per cent compared to the 12 per cent of interview-based studies. Of the self-report studies, at a threshold of 12/13, the EPDS provided the same prevalence as the combined interview-based studies, with other self-report measures finding either much higher or much lower prevalences, with the exception of the BDI (Centre for Epidemiological Studies Depression Scale [hereafter CES–D] (Radloff

1977): 18 per cent; Zung Depression Inventory (Zung 1965): 7.6 per cent; BDI (Beck *et al.* 1961): 11.6 per cent).

The two variables that accounted for about 25 per cent of the variance in the prevalence of postpartum depression were length of period assessed (17 per cent) use of self-report assessments rather than interview-based assessments (8.6 per cent variance). This meta-analysis did not cover the incidence of new PND in the puerperium.

Evans *et al.* (2001) report findings from a longitudinal cohort study of depressive symptoms during pregnancy up to 8 months postpartum in a population of 9,028 women, using the EPDS with a threshold of 12/13 both antenatally and postnatally, but also providing EPDS means. Using dichotomous variables, prevalences of antenatal depression were 11.8 per cent and 13.5 per cent at 18 and 32 weeks of pregnancy, respectively, and 9.1 per cent and 8.1 per cent at 8 weeks and 8 months postpartum, respectively. Incidence of new PND at 8 weeks postnatal without antenatal depression at 32 weeks gestation was 5.3 per cent. The majority (75 per cent) of women scored below the threshold at all four time-points and hence did not show evidence of depression throughout the perinatal period, but 1.6 per cent had probable depression throughout the antenatal and postnatal assessments. Since the authors of this study do not provide data on the number of women who remained depressed between the 2 postnatal assessment points at 8 weeks and 8 months, it is not possible to draw conclusions regarding the prevalence of chronicity of PND. A more recent report by the same group (Heron *et al.* 2004) of antenatal and postnatal anxiety and depression, which appears to use some of the same study sample (n = 8323) with slightly different prevalence results, provides estimates of numbers of women who are depressed at 8 weeks and 8 months. According to my own calculation from the numbers given, about 39 per cent of those women were depressed at both time-points (i.e. they had a score above 12 on the EPDS). It is possible that a proportion of these women suffered from a chronic depression lasting for the first eight postnatal months; however, because assessments did not take place between the 8-week and the 8-month assessments, certainty as to the number of women with such a chronic disorder is not possible.

1.1.3 Course of PND: duration, onset and recovery

Why is the course and potential chronicity of PND important? If PND were to have a likely impact on the development of infants, a more chronic disturbance of maternal mood might have a more pronounced effect on the mother–infant relationship. These considerations indicate the value of establishing the course of PND, and any factors that might contribute to vulnerability to chronicity.

Despite the importance of chronicity, not many studies have focused on this topic and the evidence is somewhat inconsistent. Several studies have shown PND to be of short duration, the mean duration of episodes being 3.3 weeks (O'Hara *et al.* 1984) and the majority lasting less than 3 months, whereas non-puerperal depression lasts longer (Cooper & Murray 1995). In contrast, there are some findings of a substantial minority of postnatal depression episodes being prolonged. Cooper *et al.* (1988) found 15 per cent of episodes to last for at least 6 months, with 4 per cent lasting almost one year. These findings were confirmed by Elliott *et al.* (2000) in a prospective study of antenatal and postnatal mood which found nearly 10 per cent of women to be depressed for at least 6 months.

Since not many studies have undertaken assessments of maternal mood over a longer period, a study of chronicity (Campbell *et al.* 1992) and its impact on infant development (Campbell & Cohn 1997) is described here in order to outline relevant issues. The authors describe the course of PND over the course of two years among a sample of low-risk depressed women and controls. Their data show depression to take a chronic course over 2 years in 30 per cent of postnatally depressed women (chronicity being defined by a continuing diagnosis of major depression; or by an intermittent diagnosis, punctuated by periods of sub-clinical depression; or by continued treatment with psychotropic medication). A further 18 per cent showed an episodic course. Although equally symptomatic at 2 months postpartum and significantly distinct from the controls, at 4 months women in the remitted group had dropped to the level of the control group, and at 6 months only the chronic sub-group continued to show statistically elevated symptom ratings, ratings which remained elevated throughout the follow-up interval. These findings indicate that some women suffer from mood disorders that reach a level of severity of diagnosis, but tend to be

brief. Those women who remain depressed show a relatively persistent level of moderate symptoms after the initial postpartum period.

1.1.4 Risks of PND

Because of the potential influence of hormones on depression in general and the dramatic hormonal changes during the puerperium, the roles of a number of hormones have been investigated in relation to PND. Findings are, however, largely inconsistent, providing no veridical evidence of the role of progesterone (Harris *et al.* 1989b; O'Hara *et al.* 1991) or of cortisol, prolactin and oestrogen (O'Hara 1997). Investigation of the role of thyroid disorder and antibodies, possibly expressed via depressive symptoms resembling PND (Harris *et al.* 1989b), has shown greater promise. Although the role of neurotransmitters has been extensively investigated regarding depression in general, there has been little focus on their role in relation to PND, presumably because neuro-chemical changes are assumed to be similar in any kind of depression.

Research has therefore increasingly concentrated on attempts to understand psycho-social factors that may contribute to or cause PND. Yet most studies have treated PND as if it were a single entity, and that is unlikely to be the case. What *is* likely, however, is that PND is a conglomeration of diverse disorders and distress – with different associated risks – which share the expression of symptoms. A review of risk factors shows that they resemble in the main those of depression in non-childbearing women, indicating that there are general factors at work. Risk for depression across women's life-cycles is discussed further in chapter 2.

The meta-analysis undertaken by O'Hara & Swain (1996), described in section 1.1.2.2, reviewed prospective studies of PND risk factors. They established that social stress in the form of relative poverty, obstetrical difficulties, stressful life events during pregnancy and the unhealthy state of a woman's marital relationship during pregnancy, pose clear risk factors for postpartum depression, with other risks being lack social support (Brugha *et al.* 1998), past history of psychopathology, and significantly dysphoric mood during pregnancy and neuroticism as assessed during pregnancy. These findings strongly suggest that there is a continuity of psychiatric

disturbance that extends back years before a woman's pregnancy and continues into her pregnancy and the postpartum period (O'Hara & Swain 1996). The nature of the links between a past psychiatric disturbance and PND appears not to have been a focus of research thus far.

The obvious major difference between PND and depression generally in women is that PND occurs when a baby is born, childbirth being likely to bring into a woman's life considerable change, to which she has to adjust. This poses the question of the extent to which it is childbearing *per se* that affects the occurrence, timing or severity of postpartum disturbance (O'Hara & Swain 1996).

Particular aspects of this question – for example whether processes involved during the delivery pose a risk of depression – have been addressed in several studies. Although a number of studies have investigated whether obstetric factors pose a risk, findings have been inconsistent and marred by methodological limitations, affording no clear answers. A recent prospective study in the UK (Warner *et al.* 1996) showed two obstetric factors in the wider sense – unplanned pregnancy and not breast-feeding at six weeks – to be independent predictors of increased risk of PND. It is possible, that these risks are proxies for other factors involved in the development of PND, and as such an expression of a woman's attitudes to pregnancy and her baby; in addition, the direction of risks can be difficult to establish, and an inability to breast-feed could be an early sign of distress or depression rather than its cause.

In contrast to these UK findings, a more recent prospective Australian study (Johnstone *et al.* 2001) obtained data on obstetric risks, personality, psychiatric history and life events. Only psycho-social factors increased the risk of PND at eight weeks postpartum, whereas none of the obstetric variables increased the risk significantly, thereby providing further support for the meta-analytic findings of O'Hara & Swain (1996).

Other potential risks associated with childbearing itself involve factors relating directly to the baby. Murray *et al.* (1996) investigated a sample of women predicted to be at high risk of developing an episode of PND and compared them to a low-risk

sample. After accounting for maternal 'blues' and maternal perceptions of the infant's temperament, infant motor functioning and neonatal irritability assessed in the early puerperium (at 8 weeks) independently increased the risk of PND by, respectively, 5 and 3.5 times.

Attitudes towards motherhood and its meaning are further factors that may predispose women to PND. This idea was explored by Sharp & Bramwell (2004) investigating expectations of motherhood and outcomes of these expectations in 205 primiparous women. Their findings support the idea that women's attitudes to mothering correspond to three orientations in line with a psychoanalytic model devised by Raphael-Leff (1991). These orientations are:

- facilitating (i.e. a woman who facilitates her baby's well-being by adapting herself to meet her baby's needs);
- regulating (i.e. a woman who regulates her baby by setting up routines to ensure predictability and to provide herself and her baby with a sense of security); and
- reciprocating (i.e. a woman who negotiates rather than facilitates or regulates, maintaining a high degree of flexibility in balancing her own and her baby's needs).

Women reporting a regulator mothering orientation were at three fold increased independent risk of above threshold postnatal depression symptom levels (EPDS >10.5). The authors suggest that a regulator orientation represents a challenge to a woman's view of herself and the sense of control she has over her life, thereby posing a vulnerability to depression.

1.1.5 Summary of PND findings

In conclusion, evidence points to PND occurring at about the same rate as other forms of depression in women, with a tendency of higher incidence shortly after delivery. The foregoing review of risk factors for PND shows that risks can be categorised into different types. Some risks pre-date the advent of childbirth, with pre-existing vulnerability interacting with other psycho-social factors, along the diathesis-stress model of Brown & Harris (1978); others appear to be related to childbirth itself. It is possible, in particular, that attitudes to pregnancy and progeny,

as well as infant factors themselves, increase the risk of depression. Most risks, with the exception of obstetric and infant factors, appear to resemble those applicable to depression in general.

Clearly, risks specific to PND, or an absence of the risks of PND that one would commonly associate with depression in general, would indicate the possibility of different processes for postnatal and non-puerperal depression. It is noteworthy that few studies have moved in this direction in an attempt to understand mechanisms. In particular, the relationship between a woman's experiences during *her* childhood and PND have yet to be thoroughly investigated.

1.2 Antenatal depression (AND)

The quality of mood during pregnancy has been viewed in different ways. In addition to such diversity of opinion, there is the widespread lay belief that pregnancy is a time of emotional well-being, and that where this is not the case a woman's emotional health will improve with the birth of her baby. The majority of studies using self-report symptom rating-scales have, however, established a pattern of consistently higher scores during pregnancy than postnatally, and the overall impression of studies using recognised diagnostic criteria is that prevalence rates of depression during pregnancy are comparable to those found after delivery (for an overview, see Green & Murray (1994). Although AND has thus been established as a potentially important issue, it has until recently remained a relatively neglected area of research *in its own right.* Instead, the tendency has been to assess antenatal mood to control for its effects on concurrently administered measures or on postnatal mood. Yet, in order to understand possible variations between distinct forms of depression, attention needs to be focused on AND as well as PND. In particular, understanding pathways from antenatal to postnatal depression may be important. Additionally, the fact that most of the relevant indices have difficulty in predicting more than 30 per cent of PND cases suggests the existence of multiple aetiologies and pathways to PND. Hence a separation of antenatal and postnatal depression and examination of their separate risks is a fist step to aid understanding.

The section following summarises the severity and prevalence of AND and its relationship to PND, before providing an overview of risk factors associated with AND.

1.2.1 Prevalence of AND

Most recent studies using self-report measures of depressive symptoms show a pattern of consistently higher scores in pregnancy than postnatally (Green & Murray 1994). This was confirmed by the large-scale study (n = 9028) of antenatal and postnatal mood using the EPDS as an outcome measure at 18 and 32 weeks gestation and 8 weeks and 8 months postpartum (Evans *et al.* 2001), already described. Depression was slightly more prevalent (13.5 per cent) in late pregnancy compared to 8 weeks postpartum (9.1 per cent), and more women moved above the threshold for depression between 18 weeks and 32 weeks of pregnancy than between late pregnancy and 8 weeks postpartum. Overall, antenatal scores were higher than postnatal scores. The distribution of scores for individual items on the EPDS did not, however, differ at the four assessment points, indicating that the nature of depressed mood does not differ before and after childbirth. In a subsequent report on part of the same study sample (n = 8323) the authors provide the following odds' ratios (ORs) for a prediction of PND from antenatal depression: persistent PND (present at both postnatal assessments at 8 weeks and 8 months) predicted from antenatal depression present at, respectively, 18 weeks and 32 weeks: 3.17 (95 per cent CI 2.29–4.37; p <.001); 6.55 (95 per cent CI 4.68–9.1; p <.001

The agreement among researchers that depression appears to be more frequent during pregnancy is a recent phenomenon (Johanson *et al.* 2000; Evans *et al.* 2001; Josefsson *et al.* 2001), whereas in the previous two decades such differences in prevalence between the antenatal and postnatal periods had not been found (see e.g. Kumar & Robson 1984; Cooper *et al.* 1988; O'Hara *et al.* 1990), although this divergence may be a product of distinct methods and the variant timings of assessment. Overall, the general impression now is of a strong association between antenatal and postnatal mood (O'Hara & Swain 1996), with ORs of about 4 between AND and PND (Gotlib *et al.* 1991; Green & Murray 1994; Johanson *et al.* 2000; Josefsson *et al.* 2001).

With such a strong likelihood of depression continuing into the postnatal period, it would be interesting to know which women are more likely to recover from AND. One attempt (Gotlib *et al.* 1991) to differentiate women who recovered from AND from those whose diagnosis extended into the postnatal period found only two antenatally assessed univariate variables to discriminate: reports of greater marital satisfaction and lower incidence of perceived stress. Neither, however, remained as a significant predictor of recovery from depression in a multivariate regression analysis.

Current evidence, then, shows that AND is at least as prevalent as PND and that both are strongly associated, one with the other; furthermore, depression during pregnancy is one of the better predictors of PND (O'Hara & Swain 1996). Investigating prevalences does not, however, indicate whether these depressive episodes in a woman's life are the same conditions, or whether they share merely a similar symptom profile, but are distinct disorders with varying risks and processes. A brief review of factors specific to AND is useful in addressing this question.

1.2.2 Risks for AND

Most studies have found an association between social problems, adverse events and personal difficulties, on the one hand, and depression in pregnancy, on the other (Watson *et al.* 1984; Martin *et al.* 1989; Zuckerman *et al.* 1989; Brockington 1996; Pajulo *et al.* 2001). Others have found associations with:

- marital conflict and lack of support (Kumar & Robson 1984; Kitamura *et al.* 1996);
- parental factors, such as low maternal care and high paternal control and paternal early loss (Kitamura *et al.* 1996);
- premorbid neuroticism (*ibid.*) and earlier psychiatric problems (Kumar & Robson 1984; O'Hara *et al.* 1990);
- poor-health behaviours (Zuckerman *et al.* 1989; Pajulo *et al.* 2001); and
- prior terminations, ambivalent attitude to the pregnancy and anxieties about the foetus (Kumar & Robson 1984; Brockington 1996).

1.2.3 Summary of AND findings

Present evidence is that antepartum depression appears to have risks similar to those of depression in women in general. These are a general vulnerability to emotional instability and psychiatric illness, stress and lack of social support. The prevalence of AND and of PND is much the same. Factors associated specifically with AND are prior terminations, ambivalence towards the pregnancy and anxieties about the foetus. Although a proportion of women show continuity from AND to PND, that this does *not* apply to *all* women, and that some of the risk factors differ, indicates the possibility of distinct aetiologies and pathways.

1.3 Does PND differ from depression in women at other times?

Having established that PND resembles general depression and AND in respect of prevalence and risk, the question remains of whether PND is distinctive in some respect. This section summarises findings of prospective controlled studies designed specifically to investigate possible differences by using control groups of non-puerperal women. An overview is provided of the main areas of investigation, which include comparisons of the symptoms, and the rate, severity and duration of depression.

1.3.1 PND rates compared to rates of non-puerperal depression

In addressing the question of whether PND differs from depression in women at other times, certain studies have involved control groups, none of which has shown differences in the rates of depression among a puerperal sample and the non-childbearing sample.

Cooper *et al.* (1988) examined 483 women antenatally and at three time-points over the first postnatal year. Rates of psychiatric disturbance were compared to a general population sample of 313 non-pregnant women obtained from a different study in a different geographical area. Psychiatric state was assessed by self-report questionnaires and the present state examination (PSE; see Wing *et al.* 1974). The

prevalence of non-psychotic psychiatric disorder antenatally, at 6 per cent, and in the first 2 postnatal assessments, at 8.7 and 8.8 per cent, did not significantly differ from rates in the general population comparator (9.9 per cent), while that of 5.2 per cent at 12 months postnatal was significantly lower. The majority of psychiatric disturbance had an onset within 3 months of delivery, and lasting for 3 months or less. Syndromal profiles and categorisation of cases were similar in both groups, consisting predominantly of depression and anxiety. Cooper *et al.* (1988) concluded that the prevalence and nature of non-psychotic psychiatric disorder in the twelve months following delivery does not appear to distinguish it from such disorders arising at other times. However, while the study group was assessed postnatally on three occasions, the comparison group was assessed only once, which does not allow comparison of the period prevalence rates of depression.

O'Hara *et al.* (1990) conclude similarly that the postpartum period is not characterised by an increased risk for non-psychotic depression. They matched 182 pregnant women with 179 of their female non-childbearing acquaintances, using self-report questionnaires and interview-based RDC criteria administered to the two groups at the same time. Neither the rate of depression at antenatal assessment (7.7 per cent childbearing *v.* 5.6 per cent non-childbearing), nor the 3–month period prevalence for PND were statistically different (10.4 per cent childbearing *v.* 7.7 per cent non-childbearing), but childbearing women showed higher levels of depressive symptoms antenatally and in the early puerperium. However, study and control group may not have been truly representative due to a selection bias. With a refusal rate of about 55 per cent in the study group, less-healthy women may have elected not to participate. The control group was chosen to resemble the study group but not childbearing women in general. Since there is a possibility that study and control group shared environments, their risk of depression may have been similar. In addition, it is uncertain how the pregnancy of a study group participant may psychologically affect an acquaintance.

Finally, Cox *et al.* (1993) conducted a study of 232 pairs of postnatal women and non-pregnant controls chosen from GP lists, matched by age, marital status and number of children. The EPDS was completed at 5 months postnatal and RDC-based diagnoses were obtained by subsequent administration of the Standardised

Psychiatric Interviews (SPI; see Goldberg *et al.* 1970) to a selection of high and low scorers. No significant difference in the point prevalence of depression at 6 months postpartum was found (postnatal women 9.1 per cent *v.* 8.2 per cent controls), nor in the 6 months period prevalence (13.8 per cent postnatal *v.* 13.4 per cent controls). The two groups differed in rate of onset of depression, with the postnatal group showing a threefold higher rate within 5 weeks of childbirth; and of duration, with a longer mean duration of depression in control (50 weeks, *SD* 64) compared to postnatal women (36 weeks, *SD* 45). This latter finding may explain why, despite higher rates of early-onset PND, the period prevalence rates of both groups are similar. This perhaps shows that there is a group which develops transient PND and that specific risks are triggered by childbirth.

In summary, the combined findings of the studies suggest that while period prevalences tend to be similar for depression in women, postnatal or not, the time shortly after the delivery of a child is one of higher vulnerability to depression.

1.3.2 Symptoms and chronicity compared to non-puerperal depression

If prevalences of puerperal and non-puerperal depression are similar, a question remains: do they differ in other respects, such as symptom distribution, severity and duration? Contrary to earlier notions that PND is a form of 'atypical' depression (Pitt 1968) recent studies have shown no great difference in the course or the form of PND to non-puerperal depression.

The study by Cooper *et al.* (1988), described above, showed no significant differences in 'cases' of psychiatric disturbance, and only minor differences in the syndromal profile at six months postnatal, from those of the comparison group.

O'Hara *et al.* (1990) similarly did not find any difference in average length of episodes between childbearing subjects and their acquaintance controls, but at the second and third trimester antenatal assessments and three weeks postpartum, childbearing women had significantly higher depression scores and reported significantly poorer social adjustment than non-childbearing women in the postnatal

period. Specifically, this applied to marital adjustment in late pregnancy and postnatally.

Little difference in reported symptoms between postnatal and non-postnatal depression was also found by Whiffen & Gotlib (1993), who suggest that PND is to be conceived of instead as an adjustment disorder. They compared an antenatal sample of childbearing depressed and non-depressed women with non-childbearing depressed and non-depressed women. The controls had been recruited by advert for a study about 'women and stress'. The childbearing depressed group had milder severity levels and better marital adjustment than the non-childbearing depressed group. Neither symptom profiles nor duration of depression differed between the two groups; with almost half the participants in each group being still depressed at 6 months follow-up. Selection bias may have been introduced by the manner of the controls' recruitment, accounting for higher levels of baseline stress and higher marital dissatisfaction among controls compared to a general population sample. This is particularly important, as the findings relating to the marital relationship in this study are at odds with those of others in this area (e.g. O'Hara *et al.* 1990).

1.3.3 Epidemiology of depression in women outside the puerperium

The epidemiological findings of PND presented thus far need to be put into the context of research into depression in women in general, which is provided in what follows. Unipolar depression is now seen to be the leading cause of disease-related disability among women in the world today. Several large population-based studies estimate the lifetime prevalence of depression in women to exceed 20 per cent (National Comorbidity Survey [NCS]: Kessler *et al.* (1993) – 21.3 per cent; WHO Study of Psychological Problems in General Health Care [WHO–PPGHC]: Maier *et al.* (1999) – 26.1 per cent). In both studies point prevalences are in the range of 13 per cent (Kuehner 2003).

There is consistent evidence world-wide that the rate of depression in women is about double that in men (Murray & Lopez 1996; Kessler *et al.* 2003). These gender differences first emerge in the age range 11–14 (Angold *et al.* 1998), the time of puberty. The reasons for this are not entirely clear, but appear to be unrelated to

biological differences, such as a change in sex hormones (Yonkers *et al.* 2000). It is likely that the higher prevalence in women than men is due to higher risk of first onset, rather than differential persistence or recurrence (Kessler 2003). When predictors of onset of episodes of major depression were investigated, a number of consistently significant factors were found, including family history, childhood adversity, various aspects of personality, social isolation and exposure to stressful life events (Kessler 1997). The role of childhood adversities for depression in women will be explored further in chapter 2 (section 2.2.2).

Taken together these findings appear to confirm that prevalence rates for women obtained from a variety of studies resemble those in the antenatal and perinatal period.

1.3.4 Evidence for the specificity of PND

Having established that in most respects PND resembles general depression, the following section provides an outline of findings in support of the case for PND as a disorder specific to childbirth in some women. Cooper & Murray (1995) prospectively investigated new-onset index episodes of PND compared to PND as a recurrence of previous depression, and found first-onset PND to run a shorter course with nearly 75 per cent lasting 1–3 months. Five-year follow-up revealed that women whose index episode of PND had been a recurrence of depression had a higher risk of further depression outside of the puerperium but not of depression after a subsequent delivery, while women whose index episodes had been de-novo PND were at greater risk of developing depression in a further postnatal period than of developing depression outside the puerperium.

A further pointer to specific and differential effects of distinct forms is provided by the finding that infants of women for whom this was the *first* episode of PND performed especially badly on object performance tasks at 9 months compared to those of postnatally depressed women with previous histories of depression (Murray 1992). Although this effect appears to be transient, the differences having disappeared by the subsequent assessment, this finding may support the notion that

some women are at risk for PND as result of specific vulnerability, which in turn may pose a specific risk to their babies.

There are a number of ways of understanding these findings. This new and transient depression may be specifically triggered by childbirth and thereafter be visited on the baby, so affecting its development. Alternatively, newly emerging maternal depression may be associated with other factors, such as maternal childhood adversities, which could (indirectly or directly) affect parenting, perhaps indicating that transgenerational processes are at play.

The argument that PND is distinctive might be strengthened if there were different risks associated. For example it has been argued anecdotally that the risk for domestic violence is increased antenatally and that this is associated with PND (Bohn & Holz 1996; Horrigan *et al.* 2000). Increasing evidence, however, points towards pregnancy not increasing the risk for violent victimisation (Gelles 1988; Jasinski & Kaufman Kantor 2001; Jasinski 2004; Bowen *et al.* 2005). What has been consistently established is a continuity domestic violence, in that women reporting victimisation in pregnancy have been victimised prior to pregnancy (Evins & Chescheir 1996; Hedin 2000; Martin *et al.* 2001; Saltzman *et al.* 2003).

If the puerperium poses a specific risk of depression for some women but not others, the question arises of what it is, in the puerperal period, that poses these specific risks and what kinds of risks are involved. The possible factors include: the adjustment to the new situation as a mother; her relationship with the new baby; the baby's personality; psycho-social aspects of the woman's life and her support; as well as her attitude to motherhood. Considering that the puerperium is a time when a woman herself becomes a care-giver, issues that are likely to arise at this time relate to her own experiences as a child and her relationship to her care-givers during childhood, areas which this study investigates.

1.4 Summary

Consistent evidence exists that there are no clear differences between postnatal and non-postnatal depression in terms of rate, pattern and symptoms, with the exception

of an increase of transient depression within the first few weeks postpartum, which possibly indicates a vulnerable period posing specific risks to susceptible women. It is likely that for some women childbirth is a time of specific risk for the development of new PND or for the recurrence of a previous one, and infants of some of these women may have a particular vulnerability. There is a risk that by its name, PND is conceived of as a homogenous entity which can be investigated and dealt with as if it had a single cause for all the women affected, with the attendant problem that research might not be conducted on understanding important mechanisms. In order to establish whether PND is a distinct form of depression, research needs to move beyond establishing epidemiological factors, and towards an understanding of individual causation and vulnerabilities.

Chapter 2

Developmental perspectives and the heterogeneity of depression

2.1 Introduction

The idea that the effects of adverse experiences may not be apparent in childhood but are seen in adult life, when they are activated by the demands of adult social roles, goes back to Freud (1961) and has been a focus of scientific enquiry since then. Over the final three decades of the twentieth century the relationship between childhood adversity and depression in later life has come in for much attention, particularly in the work by Brown & Harris (1978) whose social model of depression in women emphasised the interaction of background vulnerability factors and current stressors acting as provoking agents.

There are now widely replicated findings of an association between childhood sexual abuse (CSA) and depression in women (Fergusson & Mullen 1999; Andrews *et al.* 2002), and these have resulted in a large body of research. Although other childhood adversities, such as parental neglect and physical abuse, have been investigated, less attention has been paid to them. The exception is low-level parental care (low care), in particular by the mother, which has been shown relatively consistently to play a significant role in the development of depression (Parker *et al.* 1995; Parker 1983a). There is a possibility that such adversities are created by a shared environment or by a genetic predisposition common to parents and children, but recent studies using genetic design (e.g. Kendler & Gardner 2001) point to environmental effects after

accounting for genetics in relation to depression. In addition, a gene by environment interaction has been hypothesised so that an individual's response to environmental insults is moderated by her genetic makeup, thereby affecting the development of depression (Caspi *et al.* 2002). These strands of different research indicate that depression is now increasingly viewed as a heterogeneous disorder, the routes into which are varied and depend on interactions between care-givers, shared environment, genetics and the person as active participant in the construction of their environment. The extent to which such interaction takes place and predisposes an individual to depression stands in need of further research to help in discerning the mechanisms and pathways leading to depression and – possibly – would lead to prevention or treatment strategies in individual cases.

This chapter presents a review of the contributions childhood adversities make to depression in women. An overview of broader methodological concerns precedes the review of relevant adversities (sexual and physical abuse, parental behaviour) and their role as risk factors independent of depression. This is followed by a discussion of developmental processes and their contribution to adult psychopathology in general, leading on to a consideration of the specificity and heterogeneity of depression. Finally, I summarise current research on childhood adversity and depression in women in the perinatal period. Since the focus of this work is on women, the material presented is concerned particularly with factors affecting depression in women.

2.2 Childhood adversity

2.2.1 General methodological issues

A number of methodological considerations are relevant to the study of childhood adversity and depression.

2.2.1.1 Prospective (longitudinal) v. retrospective reporting of childhood adversity

The majority of studies examining childhood antecedents of adult psychopathology employ retrospective methods in the assessment of the nature of adversity and

trauma, and it is important to bear in mind that even studies which use prospective methods tend to include some retrospective assessments (e.g. Horwitz *et al.* 2001). Retrospective assessment of childhood adversity tends to be justified on pragmatic and ethical grounds. The latter make it impossible to establish childhood trauma prospectively without providing an intervention to the risk. Similarly, it is impossible to employ any experimental study design, such as a randomised controlled trial, to check causality of childhood adversity. Retrospective cohorts are more convenient and cost-effective to sample, with results being available faster than from longitudinal studies, where sampling takes place in childhood, thus incurring a wait while the children grow up. In addition, retrospective assessments enable the use of up-to-date assessment tools, whereas longitudinal studies rely on tools used 20–30 years previously, which are possibly outdated.

Nevertheless, retrospective studies are open to the criticism that they have introduced a number of potential biases (Maughan & Rutter 1997; Hardt & Rutter 2004), some of which relate to the recall of experiences, while others relate to conclusions that can be drawn from retrospective methods. I deal with these criticisms in order.

Because retrospective studies tend to assess individuals' recall of events that happened a considerable time ago, the quality of adult memories of lifetime psychiatric disorder (Andrews *et al.* 1999) and of childhood adversity may be variable due to forgetfulness. However, there is evidence that test–retest reliability for depression in some studies has been shown to be moderate to good (Plantes *et al.* 1988; Kendler *et al.* 1993), and the temporal stability of reports of CSA assessed at ages 18 and 21 has been shown to be high for those reporting no abuse whereas consistency of reports of having been abused was low (Fergusson *et al.* 2000). Parenting experiences, as assessed by the Parental Bonding Instrument (PBI; see Parker *et al.* 1979), have been shown to have a relatively high ten-year test–retest reliability (Wilhelm & Parker 1990).

One can recall only that of which one had been aware at the time of its occurrence, and most people are unlikely to remember much of what took place in the first 2–3 years of life (Pillemer & White 1989; Lewis 1995; Hardt & Rutter 2004). Other

factors affecting recall are the general tendency to apply meaning to memories and to assessments, so that answers may be provided in line with social desirability and expectations. It is likely that many memories, rather than being discrete, reflect – at least in part – agreement on a family narrative (Hardt & Rutter 2004). In addition, recall can be affected by current motivation and mood state; and it is possible that people with (or without) current difficulties selectively answer questions in line with the significance they attach to, and links they themselves make with, childhood experiences and current affect. In particular, it has been shown that adults without psychiatric disorder are more likely to underestimate childhood adversities (Maughan *et al.* 1995; Hardt & Rutter 2004).

A number of methods have been devised to test the extent of corroboration of adult recall of childhood memories. The ideal in any such corroboration would be the establishing of a true condition to which reports of adult memories are compared, thereby providing an estimation of the accuracy of recall. Because this ideal is usually unattainable, studies have devised methods to approximate for the 'true' condition (i.e. knowledge of the existence of childhood abuse) by using corroborative reports of other family members: Bifulco *et al.* (1997) used sisters' reports to validate recall of CSA. This study had methodological difficulties with a low response rate, and showed only half of the reports of CSA being confirmed by a sister, with higher rates for physical abuse and neglect. Significantly, corroboration was higher when both sisters were affected by the same adversity, and the lack of corroboration between discordant sibling pairs raised the possibility that the sisters were unable to differentiate between their own and their siblings' experiences. When interpreting these findings, one needs to bear in mind that sisters may have provided an indication of what they know or don't know about each other rather than of whether adult recall is accurate.

A further difficulty of validating retrospective recall is that studies have confined samples to individuals experiencing abuse that led to police/judicial action or to clinical referral (e.g. Widom's group [Widom & Shepard 1996; Widom & Morris 1997] comparing adult reports of CSA, physical abuse and neglect with those documented in childhood, obtaining relatively high false negative data). The lack of a non-abused comparison group means that, although data are obtained on false

negatives in recall, there can be no measure of false positives, nor can it be assumed that recall is the same in cases of abuse that came to official notice as for abuse that was kept secret or did not result in interventions (Hardt & Rutter 2004).

Nevertheless, studies of adults with known abuse histories (Williams 1994), of sibling reports (Bifulco *et al.* 1997; Duggan *et al.* 1998) and of genetic influences on reports of parenting (Kendler 1996) suggest that reports of major childhood adversities are reasonably robust (Hill *et al.* 2004a), particularly if assessments focus on serious adversities that are open to operationalisation, using high-quality measurements (Hardt & Rutter 2004). The main concern is that about one-third of the individuals affected do not report the occurrence of abuse, leading to underestimation of the incidence of abuse and neglect. Such under-reporting may lead to biased associations with psychopathology, in that individuals who are functioning well in adult life may forget or under-report early adversities (*ibid.*).

Finally, retrospective data have substantial limitations with respect to causal inferences (*ibid.*), in that, because of inaccuracy and biases in the recall of time sequences, most cross-sectional studies rely on between-group rather than within-individual comparisons of change. Further disadvantages are possible biases associated with refusal and the fact that comparison control groups tend to be selected at the time of the outcome under study rather than at the time of the experience of risk.

2.2.1.2 Sample selection.

How a study sample is selected is another methodological concern for the assessment of the relationship between childhood adversity and affective disorders, since sample selection may have a profound effect on prevalence estimates, and hence on levels of associations with depression. For ease of access to larger groups of participants, many studies have recruited convenience samples (e.g. students) or clinical and referred samples (clinic attenders, adult psychiatric patients), where prevalences of the risks under investigation may vary considerably. It would be unreasonable to believe that estimates of prevalences of CSA, based on such samples, reflect the prevalence of CSA in the general population (Fergusson & Mullen 1999). For

example, it has been shown that help-seeking behaviour is associated with a history of CSA (Romans *et al.* 1999; Newman *et al.* 2000; Kapur *et al.* 2004), thereby raising the risk of bias and an overestimation of associations between CSA and adult psychopathology in clinical samples. Similarly, subjects selected for diverse personal or family difficulties, or the absence thereof, may have had atypical childhood experiences and risks and are thus unrepresentative. In addition, samples drawn from particular communities and geographical areas may be unrepresentative of the general population.

2.2.1.3 Considerations of improved design of studies investigating childhood adversity and adult psychopathology

General population samples allow us to include those people in the community who have had adverse childhood experiences and/or depression but who have not, in the main, come to the attention of mental health services, thus reducing the likelihood of sample biases associated with help-seeking or referral. This provides a better test of whether there are specific associations between adversity and perinatal depression, as in this study. Nevertheless, even an approach of this kind can result in biased estimates of associations, if there are systematic sample losses, for example, because of non-participation and/or non-response in the investigation and if these are linked to the factors being investigated. Studies have shown the group of self-excluders to be more psychologically vulnerable and to be at higher risk of adverse outcomes (Cox *et al.* 1977; Murray *et al.* 2003); conversely it is possible that those who have not been exposed to the risks elect not to participate, feeling they have nothing to contribute (Haugaard & Emery 1989). These considerations suggest that the best way of assessing prevalences of childhood risks and their links to adult psychopathology is via a representative sample of the general population, using random sampling methods and assessing the extent of each sample member's exposure to the risk (Peters *et al.* 1986).

Additionally, there is a need for studies examining factors that are correlated with each other and therefore may lead to confounding, so that the interaction between the various adversities and their impact on psychiatric outcome can be better understood.

To that end, the various childhood adversities should be investigated jointly in relation to depression, an issue that is illustrated in the sections which follow.

2.2.1.4 Risks and causation

Since this study is concerned with risk factors and mechanisms of PND, I provide a brief overview of those terms prior to discussing individual risk factors.

Risk denotes the probability of an outcome in which specific factors are agents or correlates shown to precede the outcome (Kraemer *et al.* 2001). To demonstrate that an agent is a risk factor requires its examination in a sample, one which is free of the outcome of interest, and prospective follow-up to distinguish those who do from those who do not subsequently have the outcome of interest. The result divides the population into a high-risk and a low-risk group, the two comprising the total population, where the probability of the outcome in the high-risk group of subjects must be shown to be greater than the probability of the outcome in the low-risk group. A statistically significant association – which depends on the power of the study, i.e. the sample being of sufficient size and the research design–measurement being adequate to document non-random association between the agents and outcome (Kraemer & Thiemann 1987) – between proposed risk factor and outcome needs to be established, along with an indication of the risk factor's potency (Kraemer *et al.* 1997). Potency tends to be expressed as a statistical summary measure, such as an odds' or risk ratio, relative or attributable risk, kappa, phi and gamma.

Retrospective methodology in general population samples is of value in establishing associations between risk factors and outcome. However, it is to be borne in mind that more stringent criteria are required to establish causality. For a causal relationship to exist between risk factor and outcome, the risk factor needs to have occurred in time before the outcome, and be shown to be both manipulable and, when manipulated, to change the extent of the risk of the outcome.

A useful – if all but impossible to satisfy – set of ground-rules for establishing cause-and-effect relationships has been provided by Hill (1965), and includes:

- seeking natural experiments to facilitate the generation and testing of causal mechanisms;
- the testing of competing mechanisms;
- reversal of effects when causal stimulus is removed;
- multiple replication in different circumstances;
- designs to dissociate possible mechanisms;
- a dose-response relationship between causal risk factor and outcome;
- specificity of effects;
- biological plausibility; and
- a strong association between risk variable and the problem outcome.

In general, causality is difficult to demonstrate in observational studies, and randomised controlled trials – which would be the method of choice to test causality – are not appropriate for the investigation of childhood adversity for ethical and pragmatic reasons. Even if causal risk factors are established for an outcome, they do not tend to be indicative of the mechanisms or processes by which they operate.

Although the prospective design of this study avoids some of the limitations of cross-sectional studies, overall it falls short of some of the requirements for causality as outlined by Hill (1965).

2.2.2 Childhood antecedents of depression in women

2.2.2.1 CSA

Over the last twenty years, experiences of CSA in relation to the development of female adult psychopathology has become an increasing focus of attention. The proportion of the child population exposed to sexual abuse is a source of ongoing debate and controversy (Fergusson & Mullen 1999). This is related in part to variations in method employed in studies. The following factors could be considered a potential influence over prevalence figures (in this context prevalence refers to the proportion of the population reporting CSA):

- *The definition of CSA:* There is some variation in the definition employed in studies, but the majority use it to refer to sexual experiences, usually involving

genital contact, attempted or actual intercourse, or in post-pubertal girls touching of breasts. Prevalences of CSA are discussed below.

- *The period that constitutes childhood:* Some studies have used 16 as the upper age limit of a definition of childhood, whereas others have extended this to 18. A few studies have either not provided age ranges at all or even used an upper limit as low as 12 (see Fergusson & Mullen (1999) for a fuller summary). These variations in age may influence prevalence reports, although there does not appear to be a clear trend observed in any particular direction (Fergusson & Mullen 1999).
- *How the status of perpetrator is defined:* Some studies place importance on a clear age difference between perpetrator and victim. (For example Mullen *et al.* (1993) use a definition of the perpetrator being at least 5 years older).
- *Methods of measurement of CSA:* As well as varying in the quantity of questions asked to assess CSA, studies differ in their assessment methods (such as face-to-face or telephone interviews and self-completed questionnaires) and in what they treat as corroborative information (information which substantiates claims of CSA). The number of questions asked appears to influence prevalence estimates, with single-question assessments providing lower prevalence figures than those based on multiple-choice questions and items (Hibbard *et al.* 1990; Wellman 1993; Fergusson & Mullen 1999). No clear trend appears to suggest that the method of assessment (i.e. face-to-face or telephone interview or questionnaire) influences reported rates of CSA (Fergusson & Mullen 1999).
- *Problems of reliance on recall:* In general, studies of CSA have relied on adult recall of past events, with only a few studies attempting to verify such recollection or to use prospective design (Horwitz *et al.* 2001).
- *Sample selection:* The prevalence of CSA is likely to depend on the population studied, clinical populations being likely to provide higher prevalence rates than those of general community population samples. For ease of recruitment, many studies have collected convenience samples, such as college students, with the potential for biased findings if particular groups of the community (e.g. socio-economic or ethnic groups) are not included.
- *Problems of disclosure and non-response rates:* For the above reasons, the preferred way of assessing prevalences of CSA is via random sampling of a

representative cohort. However, even such attempts can lead to false prevalence estimates, since there is a risk of under-reporting and selective non-participation. Although these are issues of concern for any study, irrespective of the topic under consideration, they may be more pertinent to enquiries into CSA because of its emotional dimension. Hence, those exposed to CSA frequently decline to participate in population studies, for reasons of embarrassment and shame, or because they find such disclosure painful and upsetting. In addition, there may be a relationship between current mood and the disclosure or non-disclosure of abuse memories.

In what follows I provide an overview of currently established prevalences.

2.2.2.1.1 Reported rates of CSA in general population samples

The rate of reported CSA depends on the criteria used to define such experiences. The reported weighted average prevalence of CSA in women is about of 19 per cent (Fergusson & Mullen 1999) when a narrow definition is used (Andrews *et al.* 2002), whereas a broad definition yields a prevalence of approximately 23 per cent on the basis of the 513 studies published up to 2001 (*ibid.*). Other recent reported rates from general population studies investigating abuse prior to 16 years of age include 18 (Hill *et al.* 2001) and 16.7 per cent (Nelson *et al.* 2002), and 18 per cent in a study involving those under the age of 11 (Jaffee *et al.* 2002).

2.2.2.1.2 Characteristics of victims and perpetrators

Who are the girls suffering abuse and whence do they come?

CSA is most likely to be experienced by prepubertal or peripubertal girls; in the majority of substantiated cases the abuse begins prior to the onset of menstruation (Fergusson & Mullen 1999). More specifically, about 8 per cent of CSA cases have an onset prior to the age of 5 and a further 73 per cent prior to 16, the median age of CSA onset being between 10 and 11 (Finkelhor & Baron 1986). Approximately 10 per cent of CSA experiences will continue for more than one year (Andrews *et al.* 2002).

Children from different social strata appear to be exposed to similar risks of CSA according to Fergusson & Mullen's 1999 review. The finding that CSA is largely

unrelated to socio-economic status is contrary to the findings of most other studies of child abuse in general, where links between physical–emotional abuse risks and social class or disadvantage have been established (Dubowitz *et al.* 1987; Whipple & Webster-Stratton 1991; Connelly & Straus 1992).

Despite the lack of association with socio-economic status, there appears to be increasing evidence that CSA is linked to family difficulties and dysfunction in the form of parental conflict, separation, and the reconstitution of families to include step-parents; parental problems of adjustment, alcoholism (McLaughlin *et al.* 2000) and criminality; and measures of parent–child attachment (e.g. Mullen *et al.* 1993; Fergusson *et al.* 1996a), particularly in the form of parental rejection (McLaughlin *et al.* 2000).

There may be several explanations possible for the link found between CSA and family functionality. Children and parents share environmental and genetic influences – both of which can pose a risk of CSA – that are difficult to disentangle. For example, familial factors associated with CSA may be indicators of families likely to contain a perpetrator or which are likely to facilitate access by a perpetrator. That most incidents of CSA are perpetrated by people who are not immediate family members might indicate that these associations reflect environmental influences (such as rearing practices; relationships; life events; or social ethos or group influences (Rutter 2005)).

Fergusson *et al.* (1996b) examined the family backgrounds of children exposed to intra-familial, as distinct from extra-familial, CSA and found both groups to come from similar backgrounds. The authors conclude that general processes of family dysfunction create a social–familial ecology which puts children at risk of either intra-familial or extra-familial abuse. Hill *et al.* (2000) examined the quality of parental care and CSA in relation to affective symptoms in a population study of women. Their finding of low maternal–paternal care being associated with the risk of sexual abuse by a biologically unrelated perpetrator of a child below the age of 11, but not during early adolescence (age 11+), provides further evidence of environmental factors contributing to the risk of sexual abuse of younger children by unrelated perpetrators. It is therefore likely that young children abused by non-family

members or by strangers grow up in families where parents fail to exert appropriate levels of care. The finding of low paternal care is particularly interesting, and may suggest either that men fail, *by omission*, to protect their daughters from threats (possibly by not monitoring the environment to remove potential dangers) or that those fathers who provide insufficient care are 'deviant' in other respects and *by commission* are more likely to expose their daughters to potential abusers.

Who are the perpetrators of CSA?

Perpetrators of CSA are usually men (97.5 per cent: Fergusson & Mullen 1999) and CSA by parent figures is relatively uncommon. According to Fergusson & Mullen's 1999 data natural fathers account for about 3.3 per cent of all CSA incidents and step-fathers for 2.7 per cent, while the vast majority of abuse is perpetrated by acquaintances (47.8 per cent) and strangers (23.4 per cent). Although intra-familial abuse is less frequent, when it occurs it is more likely to be recurrent or to involve severe incidents (Fergusson *et al.* 1996a; Fergusson & Mullen 1999).

2.2.2.1.3 What are the effects of CSA?

CSA falls at the severe end of the spectrum of childhood adversity, which has generally been shown to play an important influence on mental health outcome. Other forms of adversity include environmental, such as socio-economic, and familial factors, along with other forms of abuse. It has been proposed that CSA involves four traumatogenic dynamics (Finkelhor 1990), consisting of betrayal, powerlessness, traumatised sexuality and stigmatisation. It is possible that, as a result of maladaptive coping, emotional avoidance and anxiety, attempts to control and feelings of despair emerge, which may develop into a number of adult psychopathologies, ranging from anxiety to depression and substance misuse, suicide being the most extreme (Polusny & Follette 1995).

Not all those individuals reporting CSA develop mental health or personal adjustment problems, and according to Finkelhor (1990), this applies to about 20 – 40 per cent. This suggests that CSA does not invariably lead to problems, and some people survive it unscathed, as well as suggesting the presence of some factors that may be protective or mitigate the effects of CSA (Fergusson & Mullen 1999). Factors posing risks are prolonged, severe abuse involving violence, actual or

attempted penetration and incestuous abuse by parental figures. The nature of family relationships and family support plays an important role, supportive relationships reducing risks of harmful effects. This may be due to a supportive family environment producing a stable, well-adjusted, child with greater resilience to the impact of abuse, once this occurs, and/or a more containing environment where the child can recover and establish the necessary security for normal development subsequent to the abuse (Fergusson & Mullen 1999). Among such protective factors during childhood are higher levels of paternal care, and positive peer and partnerships in adolescence (Lynskey & Fergusson 1997). Recent findings by Hill *et al.* (2001), however, suggest a differential role for relationships in adulthood, in that for women with a history of CSA positive relationships were non-protective of depression, in contrast to women with a history of neglect, for whom adult relationships did play a protective role.

2.2.2.1.4 The contribution of CSA to adult affective symptoms in women

CSA has consistently been found to be associated with depression after accounting for other adversities, and it is likely to be one of a number of causal factors of depression (Hill *et al.* 2001; Enns *et al.* 2002; Kendler *et al.* 2002; Levitan *et al.* 2003). However, most of these studies used retrospective method to ascertain cases of CSA, and it is interesting to note that one of the few prospective studies of abuse and neglect (Horwitz *et al.* 2001) comes to a different conclusion. That study followed up 641 adults with documented court cases of abuse (sexual and/or physical) and neglect and compared them to 510 controls with similarly deprived socio-economic backgrounds. The results, which combine the contributions of all forms of abuse with those of neglect, support the notion that adverse experiences affect general mental health, but do not make an independent contribution after life events have been accounted for. In line with this, the idea that adverse childhood experiences such as CSA have a general negative effect on adult mental health has been supported by several recent studies (e.g. MacMillan *et al.* 2001), particularly with respect to anxiety disorders, substance abuse, eating disorders, suicidal behaviour (Fergusson & Mullen 1999) and symptoms of borderline personality disorder (Gladstone *et al.* 1999). As the present study is concerned with depression in women, associations with affective disorders are the focus here.

A number of studies and reviews concur that a history of CSA increases the risk of depression considerably (e.g. Weiss *et al.* 1999; MacMillan *et al.* 2001). MacMillan *et al.* (2001) reporting a prevalence of 12.4 per cent of CSA among 3,678 women drawn from a general population sample, found an association between CSA and depression of an OR of 3.9. Fergusson & Mullen (1999) review seven studies examining the relationship between depressive responses and reported CSA. Criteria for the inclusion of studies were that they investigated community samples of at least 100 people using a clear definition of CSA and standardised instruments, and providing ORs. They established significant ORs in all cases, with a pooled OR of 4.3 (95 per cent CI 2.1–7.0), thereby providing support that there are moderate to strong relationships between depressive symptoms and reports of CSA, with a 'dose-response' relationship between the stringency of the definition of CSA and increasing strength of association. Those exposed to CSA do not constitute an homogeneous group, but rather vary in the extent of their exposure to abuse, with those variations being reflected in the strength of association between the extent of CSA and risks of disorder (Fergusson & Mullen 1999).

2.2.2.1.5 Is CSA an independent risk factor for depression in women?

Bearing in mind the considerations on methodological issues outlined in section 2.2.1, caution needs to be applied to the issue of causal relationships between CSA and depression. There are indications, that most of Hill's 1965 conditions for causality are met with respect to CSA. However, it has also been stated that CSA and other adverse experiences in childhood are strongly correlated, which might question causality. In order to determine the independent contribution of CSA studies therefore need to account for any other adversities, notably physical abuse and neglect. Hill *et al.* (2001) found that CSA and poor parental care (neglect and institutional care) but not physical abuse were independently associated with a DSM diagnosis of major depression in women aged 25–36. Kendler *et al.* (2002), investigating 1,942 female twins, found that CSA was associated with depression after accounting for exposure to a disturbed family environment and parental loss in childhood. These findings were confirmed by those of a study of twins by Nelson *et al.* (2002), which showed twins with past history of CSA to have higher rates of a range of adult psychiatric disorders, including DSM major depression when compared to their CSA negative co-twins, after controlling for other adversities. This

finding was not affected by zygosity, thereby providing evidence for CSA being an environmental factor in the causation of depression.

Although these studies point to CSA as a causal factor, in general the methods of studies do not permit such conclusions. This is due largely to ethical considerations which render both randomised controlled trials and prospective cohort studies investigating within-subject changes following experiences of CSA unemployable. Twin studies therefore appear to provide the best method of arriving at some answers.

2.2.2.1.6 Mechanisms in links between CSA and depression in women

What might be the mechanism by which CSA increases the risk of depression in women?

While much of the first generation of CSA research focused on establishing its presence, the second generation has attempted to delineate causal relationships between abuse and adult outcome (Briere 1988). A body of initial work focused on the importance of abuse characteristics and family variables, whereas more recently studies have turned to investigate cognitive processes. Several studies raise the potential importance of the relationship between adult psychopathology and the meaning a child attaches to abuse experiences, and the child's ability to disclose to and seek the support of others (Bulik *et al.* 2001; Wyatt & Newcombe 1990). In the presence of CSA, lack of maternal warmth (Conte & Schuerman 1988) and a poor parental and parent–child relationship (Romans *et al.* 1995) were strongly predictive of adult psychiatric difficulties, particularly when the severity and the extent of abuse were taken into consideration (Conte & Schuerman 1988), while Wyatt & Newcombe (1990) found negative outcomes to be more likely the closer the relationship between victim and perpetrator, the more severe the abuse experienced and the greater the self blame, with difficulty disclosing abuse to trusted adults.

The experience of CSA is frequently frightening, confusing, disorientating, painful, shaming, and therefore traumatic, and so likely to shape active coping strategies to regulate affect (Hill *et al.* 2001). One such strategy is the use of over-general autobiographical memories, which has been suggested to be a consequence of attempts to deal with painful memories associated with childhood trauma (Williams

& Dritschel 1988; Kuyken & Brewin 1995). In contrast, a recent study (Hill *et al.* 2004b), investigating the mechanisms of depression in a representative community sample of women, did not find an overall association between a history of depression and over-general autobiographical memories. Of the childhood adversities investigated, CSA – but not neglect, physical abuse or relationships with parents – was associated with over-general memory. Among those with major depression however, only the subgroup of women reporting CSA showed over-general memories for positive, but not for negative events, with this effect being much greater in those with later onset depression (after the age of 16), than in those with first episodes before that age. These findings indicate that adult-onset depression following CSA may be the result of a failure of a protective mechanism associated with general memories for negative events, and that specific memories of positive events may serve a protective function.

Abnormalities in the functioning of the hypothalamic–pituitary–adrenal (HPA) axis are well documented in depression. In addition to this, HPA axis and autonomic nervous system hyper-reactivity appears to be a persistent consequence of childhood abuse which may contribute to the diathesis for adulthood psychopathological conditions (Heim *et al.* 2000). Traumatic experiences in childhood (but not peripubertal–postpubertal trauma) have been consistently associated with persistent HPA axis and hippocampal structural abnormalities (Heim & Nemeroff 2001; Vythilingam *et al.* 2002). These findings support the neurobiological theory of early stress, such as CSA, leading to structural and functional changes of the HPA axis, causing a biological vulnerability to depression in adult life (Weiss *et al.* 1999).

This theory has received some confirmation from the results of a study of 32 referred women with concurrent unipolar depressive disorder, 21 of whom had a history of childhood trauma (sexual and/or physical or emotional), who were compared to 11 women with depression but with no childhood trauma and 14 healthy subjects (Vythilingam *et al.* 2002). Women with a major depressive disorder and a history of severe–repeated childhood physical and/or sexual abuse had smaller hippocampal volumes (18 per cent mean smaller volume), compared to depressed non-abused women and healthy volunteers, who did not differ in hippocampal volume. Although this study adds an important aspect and confirms similar previous findings (Stein *et*

al. 1997), there are some limitations which may require cautiousness in its interpretation. The power of this study was limited due to the small sample size, and there appeared to be a large potential overlap between a diagnosis of depression and either current or past diagnosis of post-traumatic stress disorder (PTSD) in the traumatised group, and the traumatised groups consisted of women with a potential combination of childhood traumata, so that there is uncertainty over which trauma is associated with the changes found. In addition, some of the difference may be attributable to PTSD rather than to depression. The absence of a comparison group comprising women with a history of trauma but without depression adds further limitations, since it is not possible to determine the direction of mechanisms. Vythilingam *et al.* (2002) raise the issue of correlation between the different forms of childhood adversity commonly seen, in particular between physical and sexual abuse. The various childhood traumata and adversities need to be distinguished and their relationship with each other assessed jointly.

In summary, having reviewed the key findings of the epidemiology and the effects of CSA, the evidence points overall to CSA as an independent risk factor in the development of depression in women.

2.2.2.2 Physical abuse

Physical abuse has been less intensively researched than has CSA, and there has been less focus on physical abuse in relation to depression. More importantly, physical abuse has not been particularly well defined, in contrast to the relatively clear definitions of CSA across most studies, perhaps because there is still uncertainty in some societies about whether physical punishment is deemed an acceptable form of disciplining children.

Assessment tools vary, ranging from self-completed questionnaires to interviews, and tend to combine different categories of type, frequency and severity (Drossman *et al.* 1990) with the quality of the activities of perpetrators or their results, such as choking, scalding or burns (MacMillan *et al.* 2001). Furthermore, participants in research are sometimes asked to provide subjective judgements about the severity of abuse (e.g. Fergusson & Lynskey 1997), rather than being provided with behavioural

examples, assessment in terms of which might be more objective. Like CSA, physical abuse tends to be assessed retrospectively, which adds to the definitional complication, as there are no external criteria to validate the condition; nor are there clear criteria as to what constitutes serious and less serious abuse. Moreover, some studies have used subjective appraisal of physical abuse (Fergusson *et al.* 2000), while others have tried to overcome the limitations of subjective appraisal by using external criteria such as a case having gone to court (Horwitz *et al.* 2001).

2.2.2.2.1 Prevalence of physical abuse

In line with the lack of clear criteria, prevalences of reported physical abuse vary from 13 to 21.2 per cent as reported retrospectively by adults, whereas about 10 per cent of parents admit to physical abuse (Straus & Gelles 1986). When more than 1,000 parents were asked in telephone interviews about the incidence, during the preceding year, of physical abuse of their children, 14 per cent (Straus *et al.* 1980) and, 10 years later, 10.4 per cent (Straus & Gelles 1986) reported having committed severe physical abuse.

The Ontario Health Survey, which interviewed a random population (MacMillan *et al.* 2001), reported that 21.2 per cent of 3,678 women aged 15–64 had experienced physical abuse, excluding slapping or spanking. In a study of 275 UK undergraduates aged 18–34, the prevalence of physical abuse among females was 13 per cent (Salmon & Calderbank 1996). A study of mothers in the UK aged 18–50 who had a child living at home found 18 per cent of the mothers to have experienced physical abuse (Bifulco *et al.* 1994). A study of physical punishment and maltreatment during childhood by Fergusson & Lynskey 1997) assessed reports by 18-year-olds and found prevalences, respectively, of 7.6 per cent and 3.9 per cent of regular and severe physical punishment by at least one parent. Contrary to the finding that CSA tends to be perpetrated by non-parental figures, physical abuse is usually committed by parental figures.

2.2.2.2.2 Is physical abuse associated with depression in adulthood?

Findings regarding a link between depression in adult women and a history of physical abuse in childhood are inconsistent. CSA and physical abuse are strongly associated (e.g. Mullen *et al.* (1993) found those who reported being physically

beaten as children were 6 times more likely to report sexual abuse, this increasing to nearly 22 times for victims of intercourse). Therefore, the independent contribution of each can be determined only in studies that account for other adversities. Several studies show associations between adult depression and physical abuse in childhood, which disappear when other forms of maltreatment in childhood – such as CSA and neglect (Hall *et al.* 1993; Hill *et al.* 2001; Harkness & Monroe 2002), or social and contextual factors – are considered jointly (Fergusson & Lynskey 1997). However, there are studies that have established physical abuse as an independent predictor for a range of mental health problems, including depression (MacMillan *et al.* 2001) and illness behaviours (Salmon & Calderbank 1996), when other forms of adversity were controlled for.

In summary, there is inconsistent evidence for the role of childhood physical abuse in the development of depression. It tends to be strongly correlated with other forms of abuse, and once these are accounted for the contribution of physical abuse appears to disappear.

2.2.2.3 Childhood neglect and parental care and control

It has long been suggested that aspects of the relationship between children and their parents contribute to depression in later life. Bowlby (1979) defined the tasks of parents in the rearing of their children as providing a secure affectional base, by their availability and responsiveness, as well as by encouraging their children to move progressively away from the parental base and develop social competence independently. When this is not the case, certain forms of 'pathogenic parenting' may adversely affect the ability of individuals to form positive emotional bonds with their parents and others, and can predispose to the development of neurotic attachment styles and adult psychopathology (Bowlby 1979; McLaughlin 2000).

2.2.2.3.1 Definition and prevalence of parental neglect

Tools have been developed to assess the nature of parent–child bonds, most assessments of parental style investigating dimensions of care and independence, and their opposites – parental indifference/rejection and control (Gerlsma *et al.* 1990; Parker *et al.* 1995); the latter being seen as components of parental neglect which has

been associated with a range of adult psychopathologies, in particular affective disorders.

In contrast to studies of CSA where the focus is on establishing epidemiology, the literature on parental rearing practices and neglect has placed less emphasis on establishing prevalences. One explanation for this is definitional difficulties in relation to *neglect*, although it is perhaps more a reflection of the use of many assessment tools as continuous measures, despite exceptions (e.g. Bifulco *et al.* 1994; Rodgers 1996a; Hill *et al.* 2001).

In their development of the interview-based measure of childhood experiences of care and abuse (CECA), Bifulco *et al.* (1994) report on a sample (recruited from general practitioner lists) of 395 working-class women aged 18–50 who had a child living at home. Nearly 20 per cent of this group reported parental indifference, antipathy from one of their parents and physical abuse, more than 33 per cent recalled discord in the family, 5 per cent lax parental control and 9 per cent sexual abuse.

2.2.2.3.2 The Parental Bonding Instrument (PBI)

One of the most widely used self-assessment tools of perceptions of being parented is the PBI (Parker *et al.* 1979), which assesses recollections of parental behaviour and attitudes prior to the age of 16. The original PBI separated parenting into the 2 components of *care* and *control* (over-protection) with 12 and 13 items respectively, scored on a 4-point Likert-type scale; more recently, a 3-factor model of the PBI, including abbreviated versions, using 16 or 8 items, has been advocated (Kendler 1996). Sato *et al.* (1999) and Cox *et al.* (2000) tested several models of the PBI amongst non-clinical Japanese workers and a psychiatric sample referred to a mood programme, respectively, and found Kendler's 1996 three-factor model, consisting of care, over-protection and authoritarianism to provide the best fit for the data available. However, the predictive validity of this model requires further confirmation, in particular with regard to prospective studies (Murphy *et al.* 1997).

The PBI has been shown to be largely unaffected by age, gender and social class (Parker 1989; Parker 1990), and appears to show sensitivity to cultural nuances of

parenting (Parker & Lipscombe 1979), high internal consistency, acceptable test–retest reliability over brief and prolonged time intervals (Parker *et al.* 1979; Parker 1981; Gotlib *et al.* 1988), and ratings appear to be stable over time, with a 5-year stability coefficient of 0.67–0.82 (Wilhelm & Parker 1990).

Since the PBI measures parental behaviours that usually occurred a long time ago, several questions about its psychometric properties need to be examined. One is whether current mood disorder may affect perceptions of parents, with the likelihood that current depression may bias towards particular patterns of answer. The findings of several studies concur that PBI responses are not affected by current affective symptoms (Duggan *et al.* 1998), in that they remained stable after recovery from depression or across different levels of depression (Parker 1981; Gotlib *et al.* 1988; Plantes *et al.* 1988), although Lewinsohn & Rosenbaum (1987) have shown remitted depressive and never-depressed ratings to differ from those currently depressed. The second factor, which may affect recollections of parental behaviour, is personality style. Although neuroticism and temperament do not appear to be associated with particular response patterns on the PBI (Parker 1981; Duggan *et al.* 1998), more recent evidence points towards cognitive style focusing on failures and mistakes as having the strongest association with maternal PBI scores (Parker 1993).

Since the PBI measures perceived parental behaviours, one might argue that this will not reflect *actual* parenting behaviour. Several studies have shown the PBI to have validity in measuring both perceived and actual parenting, having compared adults' recollections of their parents' behaviours to those of siblings (Parker 1981; Parker 1983b). A strong pointer towards the PBI assessing actual parental behaviour comes from a study (Parker 1981) contrasting perceptions of maternal behaviour of non-clinical subjects to their mothers' scores of their own behaviours. While the mothers generally scored themselves as more caring and less over-protective, thereby indicating a social desirability response bias, overall this study showed moderate agreements between subjects and their mothers.

In addition, studies comparing twins' ratings of parental behaviour have been conducted. Twin studies would be expected to show a differential response between dizygotic (DZ) and monozygotic (MZ) twins, reflecting genetic and environmental

differences, with higher correlations expected for MZ than for DZ twins. Results in this respect have been somewhat difficult to interpret, with one study (Parker 1986) failing to show such a differential response, whereas others (Mackinnon *et al.* 1993; Kendler 1996) found zygosity to make a difference in the predicted direction for both care and control.

2.2.2.3.3 Parental behaviours and depression

While some studies (e.g. Enns *et al.* 2002) have found parenting experiences, particularly lack of care, to be potentially causally related in a non-specific manner to a wide variety of forms of adult psychopathology, an extensive literature has reported associations between perceived low parental care, assessed using the PBI, and depression in adult life (Parker 1983a; Parker *et al.* 1995; Duggan *et al.* 1998; Enns *et al.* 2000; Hill *et al.* 2000). Findings on the role of high parental control (over-protection), also assessed with the PBI, have been less consistent (see Rodgers 1996a). Low parental care has some overlap with parental neglect, which is harder to measure using questionnaires (Hill *et al.* 2001), but interview-based studies of recalled neglect have found strong independent associations with depression (Bifulco *et al.* 1991; Hill *et al.* 2001). An association has been found by several studies of poor parental care with depression, mediated (at least in part) by poor adult love relationships (Rodgers 1996b; Hill *et al.* 2001; Gittleman *et al.* 1998). Although there is clear evidence that parental neglect and childhood sexual abuse are both linked to adult depression in women, recent findings have suggested that there are different mechanisms at work. Hill *et al.* (2001) found the quality of adult intimate relationships to affect the link between childhood neglect and adult depression, whereas this was not the case for women with CSA, for whom the risk of depression remained similar, whether their adult relationships were of good or poor quality.

2.2.2.3.4 Mechanisms of parental care and depression

What might be the mechanism for low maternal care influencing mood? Ingram & Ritter (2000) suggest that parent–child bonding may be linked to depression in adulthood through adverse effects of low levels of maternal care on the formation and activation of negative cognitive schema. The more neglect or rejection has been experienced from the mother, the more likely the person is to have self-devaluative thoughts when in a low mood. Transgenerational processes may play a part here, in

that it is possible that low maternal care is linked to and is an expression of maternal depression (Hammen 1991). Since evidence exists that maternal mood has an effect on children's development, as described in chapter 1, such considerations highlight the importance of disentangling the contribution of maternal care to maternal mood, with which this study is concerned.

In summary, parenting experiences, in particular maternal low care, have been consistently shown to be associated with adult depression in women.

2.2.3 Developmental processes and mechanisms

The relationship between childhood adversity and adult psychopathology is complex, and in order to disentangle causation the complicated nature of development, interactions between risks and outcome, and between different risks need to be considered. Principally, development is an active and dynamic process in which the meanings attributed to experiences alter their consequences, and individual pathways diverge in both their origins and endings (Cicchetti & Rogosch 1999). Our understanding of the mechanisms involved is still somewhat limited, and it remains a striking feature of research of psycho-social stress and adversity that there is such diversity in people's responses to similar stimuli. It has long been recognised that single risk factors lead to diverse disorders and the same disease endpoint may be arrived at via varied causal pathways (Rutter & Sroufe 2000). For example, it has been clearly demonstrated that there are genetic factors that predispose individuals to depression in adulthood, and similarly there is evidence that childhood adversity (e.g. CSA, disturbed family environment, parental loss), personality factors (neuroticism, low self-esteem, early onset anxiety), stressful life events and lifetime trauma, the quality of relationships (e.g. history of divorce, lack of social support), history of low educational attainment and past history of depression pose risks (Van Os *et al.* 2001; Kendler *et al.* 2002). Why these are risks for some people but not for all is the subject of an ongoing debate.

In most cases, psychiatric disorders are not caused by single genes, but are multifactorial, with the likelihood of genetic vulnerability being moderated by or mediated via environmental and other factors. Genetic influences are probabilistic,

not deterministic, and it is likely that overall environmental and genetic factors are of equal importance (Rutter & Sroufe 2000). However, additional 'third forces', beyond genetic and environmental influences, may play a part in development (Molenaar *et al.* 1993).

Because of the probabilistic nature of genes in relation to psychiatric disorder, it is particularly important to identify the conditions that enable the expression and modification of the underlying genetic condition, such as the contribution of familial, environmental or stress factors.

Development is not a linear process, there being a number of factors which contribute to its nature and form. It is important to recognise that individual change is due to internal processes – the structure of the brain, alteration of neuroendocrine functioning, modification of affect regulation or changes in the cognitive set or processing, altered patterns of interpersonal interaction are contenders – which then interact with external processes. A key feature of the developmental psychopathology research perspective over the last 2 decades has been the reciprocal nature of individual–environment interaction (e.g. Plomin & Daniels 1987), which recognises that the individual actively selects and shapes his or her environment, thus evoking responses from others and also contributing to the selection of and exposure to risks, while at the same time the environment itself affects the individual. This individual-environment interplay characteristically comprises complex chains of causation and events, and it is likely that such evocative responses are partly the result of a gene–environment interaction (i.e. genetic factors have an impact on behaviours that shape or select environments and, thereby, influence the likelihood of experiencing stress or adversity (Rutter 2005)). Take genetic risk of depression, life events and personality factors (neuroticism) as an example: there is evidence that genetic factors have a direct influence on depression, as do other factors, such as stressful life events and neuroticism. However, genetic factors contribute to the ways in which people deal with their environments, and evidence suggests that 10–15 per cent of genetic effects on liability to depression are mediated by a mechanism whereby individuals select themselves into high-risk environments – gene–environment correlation (Kendler & Karkowski Shuman 1997; Van Os *et al.* 2001). Recent findings suggest that such genetic effects are mediated by personality

characteristics, in that neuroticism has an indirect effect on the likelihood of experiencing stressful life events (Van Os *et al.* 2001).

In summary, it is possible to understand the complex links between early adaptation and later disorder only through an appreciation of the nature of the developmental process itself (Rutter & Sroufe 2000). It is important to delineate continuities and discontinuities between normality and the pathology of disorder. For example, the phenomenon of depression clearly spans normality and disorder in a dimensional fashion, while it is possible that bipolar affective psychoses are discontinuous with normal variations of mood (Perris 1992).

2.3 Heterogeneity of depression and its mechanisms

2.3.1 Evidence of heterogeneity

Having established the complex nature of contributions to normal and pathological development, in this section I discuss the issue of heterogeneity before proceeding to outline some possible mechanisms and processes involved in the development of depression.

For a considerable time it was tacitly assumed that all forms of depression could be treated in the same way, with the exceptions, perhaps, of bipolar and severe melancholic disorder. More recent findings, however, have queried this notion, suggesting that different processes are in operation depending on the specific features of the depression, thereby raising the possibility that depression is not a unitary but rather a heterogeneous disorder. Evidence supporting the notion of heterogeneity comes from a number of different strands of research, which show an individual's propensity to suffer an episode of major depression to be affected by, *inter alia*, life events (Brown & Harris 1978; Kessler 1997), genetic risk (McGuffin *et al.* 1996; Sullivan *et al.* 2000), relationships (Hill *et al.* 2001), history of childhood adversity in the form of trauma (Fergusson & Mullen 1999; Kendler *et al.* 2002) and parental loss (Tennant 1988; Brown & Harris 1993), patterns of co-morbidity (Kessler *et al.* 1996), the number of previous episodes (Kendler *et al.* 2001), differences in the time of onset (Jaffee *et al.* 2002) and depressive-style attributions which may play a role

in the condition's persistence or recurrence (Teasdale & Barnard 1993). In addition, endogenous depression has been associated with severe childhood adversity, particularly CSA (Harkness & Monroe 2002). It is possible that these risks result in a rather general effect, expressed in a commonality of symptoms, but each one with specific underlying processes leading to depression.

In contrast to the large body of research that has established the various risks outlined above, our understanding of the underlying mechanisms is to date still limited. I attempt to draw together relevant findings in what follows.

2.3.2 Mechanisms

Multiple reports concur that there appear to be differences in the role of social adversity in recurrent *v.* first-onset depression with depressive episode onsets becoming more autonomous and progressively less linked to environmental adversity – the kindling hypothesis (Post 1992). Further examination of this phenomenon has shown episodes of depression in individuals at higher genetic risk to be less associated with life events and, in the absence of prior depressive episodes, to frequently experience depressive episodes without major environmental stressors (Kendler *et al.* 2001). This points to the possibility that depressive episodes with little provocation can be arrived at via distinct pathways: one of numerous prior depressive episodes, perhaps driven by multiple adversities, and the other by high genetic risk (Kendler *et al.* 2001).

A history of childhood trauma may be a further source of heterogeneity, with CSA likely to be an independent risk factor for major depression in women (Fergusson & Mullen 1999; Kendler *et al.* 2002; Nelson *et al.* 2002) which can have an additive effect. There is, moreover, a suggestion that CSA has a unique, and substantial impact on a wide range of risk factors for major depression that cannot be accounted for by other associated risks, such as disturbed family environment, parental loss or genetic disposition (Kendler *et al.* 2002). Recent studies attempting to examine depression in sub-groups with childhood trauma have provided further indications of the value of conceiving of depression in terms of heterogeneity. Nemeroff *et al.* (2003), for example, examined treatment response to psychotherapy and/or

antidepressant medication in a large group of depressed patients and found it to differ dramatically when sub-groups were examined. In particular, it was apparent that for those depressed patients who had a history of early childhood trauma (early parental loss, physical/sexual abuse or neglect) psychotherapy – Cognitive Behavioural Analysis System of Psychotherapy (CBASP), a combination of cognitive therapy and interpersonal therapy – alone was superior to antidepressant mono-therapy, and when both were provided together there was no improvement in the treatments' efficacy. This effect was particularly prominent for those with chronic forms of depression and parental loss. These findings indicate not only a differential treatment response among sub-groups of depressed persons, with and without childhood adversity but, more importantly, differences between the various forms of childhood trauma. In this connection, it is interesting that an apparently relational psychotherapy has a stronger effect on those with parental loss.

Further evidence that women differ in their propensity to develop depression depending on the type of childhood adversity experienced comes from Hill *et al.* (2001). It had been established that the quality of intimate relationships is generally important in relation to the development of depression; yet Hill *et al.* (2001) found good quality adult relationships to be protective against the risk of adult depression in women who had experienced childhood neglect, whereas CSA was associated with a risk of depression that was unaltered by the quality of adult relationships. This indicates different pathways in which adult interpersonal processes, and especially attachment and cognitive processes, may play different roles. In particular, it is possible that certain strategies for coping with traumatic memories and associated affects are effective in childhood but ineffectual in the context of the demands of adult intimacy and sexuality, in that a relationship being sexual rather than the quality of the relationship may act as a precipitant. Taken together, the findings of these two studies (Nemeroff *et al.* 2003; Hill *et al.* 2001) indicate differences in the route to depression for those with a history of, on the one hand, neglect/loss and, on the other, sexual trauma in childhood.

The juvenile *v.* adult onset depression distinction provides another pointer towards heterogeneity and a differential aspect to the role of CSA. Jaffee *et al.* (2002) found juvenile onset to differ from adult onset in several key respects: a family history of

increased antisocial behaviour and higher indices of family instability, but not of CSA. The adult onset group, in contrast, showed an increased rate of CSA and residence changes, but not any of the other indices of family instability.

Following a similar line of thought, Hill *et al.* (2004a) contrasted factors associated with whether depression in women had started before or after the age of 16, particularly focusing on the type of sexual abuse experienced in childhood. They found juvenile onset adult depression to be associated with co-morbid childhood psychopathology and peer problems, poor parental care and penetrative childhood sexual abuse. In contrast, adult onset depression was characterised by non-penetrative contact CSA and, to a lesser extent, poor parental care. The authors hypothesise that the risk of adult onset depression being associated with contact CSA arises because individuals are vulnerable when they start to have romantic adult relationships involving sexual intimacy. In addition, it is likely that the juvenile onset group, which appeared to be characterised by much higher pervasive psychopathology, was less able to resort to protective cognitive processes, thereby resulting in earlier depression.

The role of the HPA axis in the development of depression and in relation to childhood trauma was outlined earlier (section 2.2.2.1.6). The findings described give rise to the possibility that early trauma alters the HPA axis and makes structural changes to the brain, causing vulnerability to later depression. However, current evidence does not help in the identification of the direction of effects. It may therefore also be likely that an underlying vulnerability of the HPA axis predisposes to maltreatment via genetic links or otherwise, and consequently leads to depression. In addition, it is possible that functional polymorphism on genes (e.g. the MAOA gene) moderates the impact of early childhood maltreatment on the development of a variety of adult psychopathologies in line with preliminarily findings by Caspi *et al.* (2002) in relation to the development of antisocial behaviour in males.

2.3.3 Implications of pathway research in the understanding of the heterogeneity of depression

This background sets the scene for my consideration of antenatal and postnatal depression. There is an indication that depression is a heterogeneous disorder. That heterogeneity is the result of the interaction of diverse risks which, although following different pathways, result in a shared symptom or syndrome profile. Furthermore, there is evidence that different childhood antecedents lead to depression via different mechanisms, following direct or indirect pathways. The trajectory of depression is further complicated by the possibility that some people are at risk of developing further episodes of adult depression if they have had prior experience of depression, while others may have silent risks, i.e. risks which were not evidenced earlier in life. Whether depression occurs prior to, or soon after, childbirth may form part of the heterogeneity jigsaw. Thus, the study of heterogeneity and of such developmental pathways, and the risks posed by previous episodes of depression, provides a fruitful strategy in researching postnatal depression.

At this point I return to PND, to discuss it with reference to the established contributions of childhood adversity.

2.4 Childhood antecedents of antenatal and postnatal depression

Few studies have assessed possible childhood sources of vulnerability to depression in the perinatal period. Although one would wish to give priority to representative studies, because of the paucity of research in that area, studies are summarised here if they are relevant to the investigation of childhood adversity and perinatal mood, irrespective of methodology employed.

2.4.1 Experiences of parental behaviour in childhood

Psychoanalytic theory emphasises the importance of an individual girl's relationship with her mother, and there is some support for one aspect of this from Brown &

Harris 1978 finding of the impact of childhood separation from the mother on the development of depression in adult women. In relation to PND there have been inconsistent findings regarding early separation from the mother, with some (e.g. Frommer & O'Shea 1973) offering evidence of association, while others (e.g. Paykel *et al.* 1980; Kumar & Robson 1984) were unable to discern such a link. The latter study, however, did establish an association between PND and early separation from father. Separation from parents in childhood may be an indicator of childhood neglect, and the development and use of instruments such as the PBI in PND research can contribute to a deeper understanding of diverse aspects of the quality of early relationships with parents in relation to PND, particularly with regard to the dimensions of care and control. Some of the studies which have investigated parental care and control are described in what follows.

In a prospective study of a representative sample of 730 women attending clinics for antenatal care (Gotlib *et al.* 1991), low levels of both maternal and paternal care, as assessed antenatally, were independently associated with the onset of postpartum depression, as assessed by the BDI (Beck *et al.* 1961) within the month following delivery and as confirmed by RDC (Spitzer *et al.* 1978) diagnosis.

Evidence of a distinct role of maternal care and control in the course of adult depression is provided by Gotlib *et al.* (1988), who administered the PBI and the maternal care and control scales of the BDI to 201 women 3 days after delivery, matching those with depression ($n = 25$) to controls and following them up 2–4 years later using the same instruments. This yielded three groups:

- women with no depression at either time point ($n = 20$);
- women depressed at both times ($n = 8$); and
- women whose PND had remitted at the second assessment ($n = 11$).

Depressed and remitted women differed at neither assessment with regard to perceived level of maternal control, but both groups differed from the women who were not depressed at either assessment in rating their own mothers as more controlling. In contrast, both depressed and remitted women differed at both assessments with regard to the reported level of maternal care, in that remitted

women resembled those who were not depressed, whereas depressed women scored their own mothers as significantly less caring. The level of perceived caring measured in the early postnatal period was predictive of the level of depression at 30 months, and thus differentiated those depressed from those remitted, both of which groups had reported high levels of control. The authors postulate that the fact that both the depressed and the remitted group showed high levels of perceived control (over-protection) may mean that this is related to vulnerability to 'general' or 'diffuse' psychological symptoms, particularly during periods of stress such as that associated with childbirth. Some caution, however, seems necessary when interpreting these results. This study included a low number of participants, which resulted in small cell sizes, thereby raising issues of power. Postnatal assessments were conducted very early in the puerperium, at day three, a time when there is a risk of confounding by other factors, such as possible symptoms of maternal 'blues' and physical complaints. The latter point is particularly important, as Gotlib *et al.* (1988) used the BDI to establish depression. As was outlined earlier (section 1.1.1.1), concerns have been raised about the postnatal use of the BDI, concerns which would apply particularly to the use of the BDI at such an early stage in the puerperium. In addition, the participants in this study exhibited relatively mild levels of depression.

Further evidence of the role of perceived maternal and paternal care in the prediction of onset of PND is provided by the same research group (Gotlib *et al.* 1991), who contrasted this with the lack of prediction of recovery in the postpartum period from depression during pregnancy.

An indication that different factors play a role in relation to the timing of the onset of PND is provided by Matthey *et al.* (2000). Their study, which did not use a representative community sample but selected participants to be representative of defence style, found maternal over-protection, or control, as assessed by the PBI, to play a role in the development of PND at 6 weeks, while at 4 months postnatal low maternal care and high maternal control were significantly associated. However, when antenatal mood was included in regression analyses, the effects of maternal behaviours disappeared.

A representative study (Kitamura *et al.* 1996) of antenatal depression among 1,289 women in Japan, which used an abbreviated 16-item version of the PBI and Zung's self-rating depression scale (SDS) (Zung 1965) to assess depression, found independent associations with maternal care and paternal control scores, and with early paternal loss. This study appears to have had a problem involving relatively high levels of missing data, especially for the PBI variables.

When considering the relationship between perceptions of parenting and the development of PND one may wish to consider whether childbearing has an effect on how parents are perceived. Whiffen & Gotlib (1993) compared 2 groups of childbearing (depressed and non-depressed) and 2 of non-childbearing (depressed and non-depressed) women in relation to a number of psychosocial variables, among which were their perceptions of parental behaviour during childhood, by using the PBI as assessed antenatally. According to the study's published means consistent differences are shown between the childbearing and non-childbearing groups, irrespective of whether or not they were depressed. The childbearing groups had higher maternal and paternal care scores and lower maternal over-protectiveness scores than did the non-childbearing women. These findings might indicate that pregnant women perceive their parents more positively than do non-pregnant women, and so as less over-protective–controlling and more caring. This gives rise to the possibility that particular processes at work during pregnancy are preparing women to their forthcoming motherhood, part of which involves a 're-evaluation' of their own childhood experiences of having been parented. This is somewhat at odds with the finding that perceptions of parents are stable over time (Wilhelm and Parker 1990b), which however did not take childbearing status into account. Hence, there is a possibility that childbearing in itself has an effect on perceptions of parenting. It is worth entering the caveat: in Whiffen and Gotlib's 1993 study the non-childbearing participants were recruited via advertisements. It is therefore possible that the difference in ratings between childbearing and non-childbearing women is an artefact of self-selection into the study by non-childbearing women with a bias towards more negative views of their parents.

The combined conclusion of these studies supports the proposal that perinatal depression is associated with childhood experiences of parents' behaviours, although interpretation of these studies is somewhat constrained by the differing measures used to assess depression and by the absence of measurements of other childhood adversities.

2.4.2 CSA and physical abuse

Despite the clear association between CSA and depression in women, there is surprisingly little research into CSA in relation to antenatal or postnatal depression. Physical abuse has similarly been ignored. In addition, there are no studies investigating the prevalence of CSA among pregnant women. If women with CSA experiences were to become pregnant at the same rate as women with no such experiences, one could expect the prevalence of CSA to be the same among pregnant women.

An indication that the prevalence of CSA in pregnant women is similar to that expected in non-pregnant general population samples is provided by Senior *et al.* (2005) who investigated the relationship between maternal eating disorder and early childhood experiences. In this large general population sample of pregnant women (n=10,641) 18.2 per cent of women disclosed early sexual contact abuse before the age of 16.

Of the few studies looking at CSA in relation to pregnancy, most concentrate on the effects of CSA on the outcome of delivery. This follows the hypothesis that early traumatic experiences activate corticotrophin-releasing hormone gene expression in the brain, and therefore predispose to elevated CRH gene expression in the placenta, thereby rendering those exposed to early abuse more susceptible to stress with possible adverse outcomes in the baby. Two such studies (Benedict *et al.* 1999; Grimstad & Schei 1999) did not find an association between CSA and pregnancy outcome or delivery variables. These studies did not investigate PND, but Benedict *et al.* (1999) found CSA to be significantly associated with higher levels of antenatal depressive symptoms, negative life events, and physical and verbal abuse prior to and during pregnancy.

Several literature searches failed to reveal prospective studies of a link between CSA and PND. Although anecdotal reports of such a link exist (e.g. Buist 1998; Smith 1998), the evidence for this is sparse, and the literature consists merely of a few case reports in the nursing/midwifery literature (e.g. Smith 1998) and research by Buist's group in Australia (Buist 1998; Buist & Janson 2001) on admitted women with PND. Buist & Janson (2001) investigated a clinical sample of 56 depressed women, admitted to a mother and baby unit, and their partners over a 3-year period regarding the impact which a history of sexual abuse has on parenting capacity in respect of mother and child outcome. At entry to the study, all women had either a DSM-III-R diagnosis of depression or adjustment disorder, and scored above 15 on the Hamilton Depression Rating Scale (HDRS; see Hamilton 1960) and above 23 on the BDI (Beck *et al.* 1961). Social support and relationship with partner, parenting of the child, mother–infant interaction and the child's behaviour were also assessed. Contact sexual abuse was established using questions from the Otago Women's Health Survey (Martin *et al.* 1993). Physical and emotional abuse was assessed using unstandardised assessments and grouped together. The initial stage of the study found an abuse history of any kind to have a potentially deleterious effect, particularly on mother–infant interaction, higher BDI scores and on longer admissions. At three years, following some attrition from the study, those with a history of CSA ($n = 23$) were compared to those without ($n = 22$). Women with CSA had higher depression and anxiety scores than non-abused women, had made less improvement over time and showed more life stresses. There was no difference in the children's cognitive scores, although these were low in both groups. The abused women's partners rated themselves as more comforting and their children as more disturbed. There was an indication that abused women had an impaired mother–infant interaction in the postpartum period, which was predictive of a higher HRDS (Hamilton 1960) score three years later.

In conclusion, there is some indication that a history of CSA in already postnatally depressed women negatively affects mother–infant interaction and infants' behaviours. Since there a no studies in which CSA and physical abuse are assessed together with parental care and control, it is impossible at present to draw conclusions about independent contributions made by these potential risks.

2.4.3 Current understanding of childhood antecedents of PND: implications for heterogeneity

A thorough literature search failed to find any studies comparing childhood adversities associated with PND and those associated with depression at other time points.

There are nevertheless good reasons to think that particular childhood adversities have specific relevance to vulnerability to depression following childbirth. A history of CSA, physical abuse and neglect may lead to an activation of issues in relation to parenting and changes to the sense of self at the point of entry to motherhood. Specifically, CSA may lead to activation of traumatic memories with regard to feelings of helplessness in a medical setting, particularly one involving a sense of being invaded by others. A history of physical abuse may put women in a position that mirrors the vulnerability of experienced trauma. Early experiences of neglect (i.e. low care and high control) may be activated by the role changes inherent in becoming a mother and having to take responsibility for a baby. It is hypothesised, therefore, that at the time of childbirth and entry into parenthood, previous adaptive strategies and defences developed to deal with adverse early experiences become overwhelmed and cease working, resulting in a path to depression.

2.5 Purpose and hypotheses

For the purpose of examining whether there are distinctive childhood antecedents of PND, three hypotheses are considered: -

(1) Childhood adversities (such as parental low care and high control; childhood sexual abuse (CSA) and physical abuse) are predictive of PND.

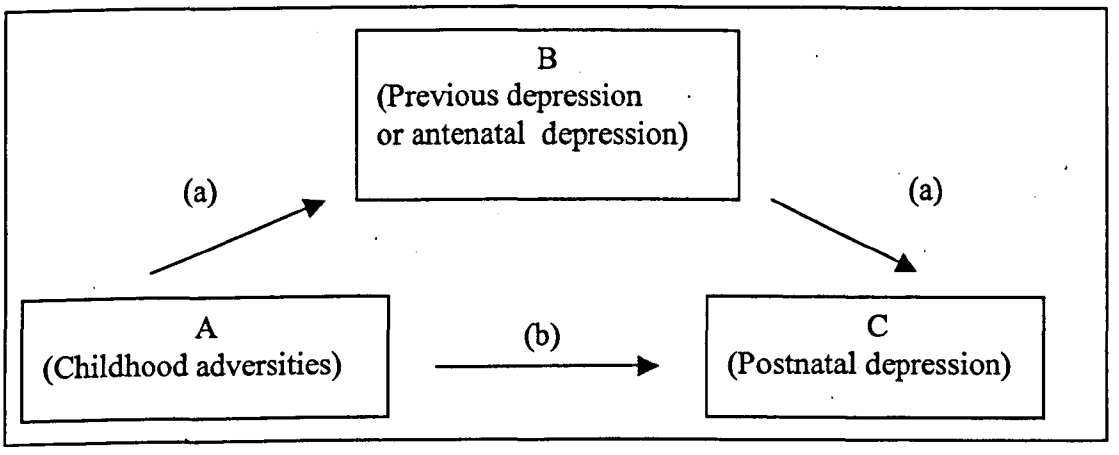
(2) Childhood predictors of PND are distinctive and differ from both previous depression and antenatal depression.

(3) There will be independent associations of childhood adversities with PND that are not mediated via a history of depression (PrevD) or antenatal depression (AND) (see figure 2.1, where (a) signifies mediation and (b) independent effects; for a discussion of the mediation model, see 3.6.2).

(4) Risks from adverse childhood experiences for newly developing PND will differ from both those of depression in the antenatal period and those occurring prior to pregnancy.

Analyses addressing these hypotheses compare and contrast transient and persistent PND.

Figure 2.5 Representation of hypothesis 2



Chapter 3

Methods

3.1 Introduction

Bearing in mind the difficulties outlined in chapters 1 and 2 regarding the design of studies into depression using assessments of childhood adversities, this study attempted to minimise potential flaws. In order to investigate the hypotheses described at the end of chapter 2, this study used a prospective design with retrospective assessments of childhood adversities. A representative cohort of pregnant women was recruited antenatally and followed up longitudinally over the first three months of the postnatal period in order to establish the outcome under investigation, i.e. PND. For the assessment of childhood adversity well established tools were used, such as the PBI for recollections of parenting; a CSA questionnaire devised by Mullen *et al.* (1993); physical abuse using questions previously designed for use as part of a study in a neighbouring district (Hill *et al.* 2001). AND and PND were assessed with the most commonly used screening tool, the EPDS. The potential childhood adversities (CSA, parental care and control, and physical abuse) were established antenatally and examined in logistic regression models jointly to predict PND; mediation by AND and any independent contribution by each of the childhood risks. The sections that follow detail the exact methods employed.

3.2 Main study

The initial tasks were to estimate the number of women who needed to be recruited, by performing power calculations (see section 3.7). Next, links were established with clinicians and management at the Liverpool Women's Hospital (LWH), whose patients were to be recruited for the study, to obtain permission to conduct the study

and to enlist the support of community midwives (CMWs), who were due to recruit participants. At the same time, administrative procedures for the running of the study were devised and piloted over the first five months of the study and mechanisms implemented for feedback on any recruitment problems. Regular reviews of recruitment rates and procedures throughout the duration of the study ensured that the administration was responsive to problems identified while attempting to minimise any biases introduced.

3.2.1 Study sample

The study population was comprised of nulliparous pregnant women, who:

- booked into the antenatal clinic at LWH;
- were due to receive antenatal care from community midwives at LWH;
- lived within Liverpool city boundaries;
- were under the care of a Liverpool general practitioner (GP).

3.2.2 Procedures

3.2.2.1 Recruitment

Permission had not been granted by the local Research Ethics Committee (REC) for a direct approach to potential study participants by the author; this had to be initiated instead by the usual health care team. Ascertainment of eligible women, therefore, was initially via the CMWs of three community groups who agreed to recruit study participants in GP-based community clinics after 24 weeks gestation. Each group comprised 4–8 CMWs, who covered 8–11 GP surgeries where weekly community antenatal clinics were held and who were regularly drafted in to cover hospital units. Their workload, it was clear, could lead to absences from their community clinics, which would have to be covered by colleagues, some of whom were not part of the usual team, so that there was a risk that involvement with study participants would be discontinuous. Early on, attempts were made to address these issues to ensure a smooth running of the study. These are outlined below.

Eligible women were identified on the hospital database and CMWs, furnished with the names of women fulfilling the inclusion criteria, were asked to approach those women in clinics. Initially the study was introduced to participants by CMWs with a

view to obtaining their consent at the next weekly appointment. This did not work well as different CMWs were involved in this process and some women were lost to the study because they had an unknown consent status. In order to minimise losses to recruitment, information was provided and consent acquired (or not) at one and the same clinic appointment. Within the limitations of the consenting process, the extent of refusal during the first five months appeared to be similar to that for the later stages of the study, and the data are therefore combined in the main results.

Close involvement of the CMWs in the administration of the study was important for several reasons. The study covered difficult topics and CMWs had a safeguarding role in that they might have to deal with women becoming distressed as a result. (A procedure for managing distress was devised for that purpose (see Appendix C) but was not required; there was no indication that any woman became distressed during the course of the study). In addition, it was felt that CMW involvement would help with response rates, and they were initially asked to hand out questionnaires for recruited participants to complete. This did not work well, for much the same reasons as outlined above: it became clear that in some clinics CMWs did not receive mail addressed to them despite attempts at providing them with a box file so that the study materials could be kept in the surgery and the approaches made to all GP surgery receptionists for their support. Following reviews, the system was changed so that midwives passed participants' consent forms to the author and questionnaires were sent direct to participants by the author.

Once recruitment and administrative procedures had been refined, the study was expanded to two further CMW groups for another eight months. This worked well, but did not yield an adequate number of participants. Because of these concerns, recruitment was transferred to the antenatal booking clinic at LWH as soon as other research studies recruiting participants there had terminated. Now, participants were recruited in the early stages of pregnancy by any CMW covering the clinic. All referrals were reviewed prior to clinics taking place and case notes of potentially eligible women identified. As fifty-seven midwives were covering this clinic, regular meetings with the antenatal booking clinic manager ensured adequate recruitment procedures and rates.

Moving recruitment to the antenatal booking clinic increased the rate of recruitment during the last seven months of the study to an acceptable level. Because patients were now recruited at an earlier gestational stage than in community clinics, attempts were made to ensure the simultaneous completion of the antenatal (time 1) questionnaire. This included telephone contact with all participants to check addresses and the adequacy of the information they had received, and to ensure that women were not lost between recruitment and receipt of the first questionnaire. Permission from the Liverpool Research Ethics Committee (LREC) to approach women by telephone was sought and granted.

Information provided by CMWs at the recruitment stage included the participant's name and address; expected date of delivery; and current gestation and consent status with signed consent form.

3.2.2.2 Consent

Eligible women were given verbal and written information about the study by CMWs and asked to provide written consent. At the consenting stage, all women approached were asked to provide some basic demographic data (age, occupation, postcode, and at a later stage telephone numbers) for later comparison of respondents with refusers and non-respondents (see 4.1).

3.2.2.3 Questionnaires

Once completed consent forms had been received, women were sent an antenatal (time 1) questionnaire (see 3.3.1; Appendix A) after 23 weeks gestation. Those who returned the antenatal questionnaire (and a small sub-group of those who did not) were sent a 6-week postnatal (time 2) questionnaire after 5 weeks post-partum (see 3.3.2; Appendix B). The 12-week postnatal (time 3) questionnaire (3.3.3; Appendix B) was sent to those participants, after 11 weeks post-partum, who had returned the antenatal (time 1) questionnaire. However, as this measure was introduced 5 months into the study, the first 140 women recruited did not receive it.

Non-respondents were sent reminders together with further questionnaires after two weeks of non-response to any of the questionnaires. All women were regularly asked to provide details of changes of address.

3.2.2.4 Interviews

In view of the distinctive demographic characteristics of the population in Liverpool, diagnostic interviews to establish major depression were carried out, to establish the appropriate cut-off of the Edinburgh Postnatal Depression Scale EPDS in this socio-economically deprived sample. For this, a random sub-sample of 10 per cent of those women who had returned their 6-week postnatal questionnaire was chosen for interviews. Returned EPDS, which were part of the 6-week postnatal (time 2) questionnaire, were scored by a research secretary, and women with a score above 10 and a control group of those scoring below 10 were drawn using the random selection facility of 'select cases' in the Statistical Package for the Social Sciences (SPSS inc. 2001). The women selected were invited by telephone to participate in a postnatal interview of DSM-IV diagnosis of depression (SCID; First *et al.* 1994) within 4 weeks of completion of the 6-week postnatal (time 2) questionnaire. Those consenting were visited at home by one of two interviewers (PB and the author), both of whom are psychiatrists. Interviewers were blind to the women's PND (EPDS) scores and rated the interviews immediately after they had taken place. Regular meetings between the interviewers and the supervisor of the study (JH) were held to iron out any diagnostic uncertainties. Inter-rater reliability tests of 10 interviews showed a 100 per cent concordance of diagnoses (3.5.1.1 provides further details of the results of SCID interviews).

3.2.2.5 Study administration

Regular attempts were made to ascertain the sample to be recruited from the database at LWH to aid recruitment and for later comparison of the recruited cohort. In the first stages of the study, when recruitment took place in community clinics, lists of eligible women were prepared for each clinic location and sent to midwives on a monthly basis.

Women's delivery dates, knowledge of which was important to ensure that postnatal questionnaires could be sent out within the time frame of the study, were not available in a systematic fashion, so that LWH data sets had to be searched weekly (both manually and electronically) for each individual participant to establish actual dates of deliveries and changes of address. GP surgeries were contacted if there was uncertainty over a participant's address.

Information was sought to establish pregnancy losses or potential deaths of babies to avoid unnecessary grief through further contact with the study. For this purpose, monthly data were provided by a LWH department on all late terminations, miscarriages and infant deaths. Another department provided information on early miscarriages, which was particularly relevant towards the later stages of the study, when women were recruited at an early gestation, giving rise to the possibility of early miscarriages in the time between booking and the posting of the antenatal (time 1) questionnaire. Information on pregnancy loss was regularly compared with the study database and the women who had suffered loss were no longer included.

3.2.2.6 Preparatory meetings with management and staff

Prior to and following approval of the study by the local REC and the LWH research and development (R&D) committee, planning and review meetings were held with the clinical director of obstetrics and the manager of the CMWs, both of whom were very supportive of the study. Several initial information meetings were held with CMWs, and once the study was underway regular monthly meetings with each CMW group recruiting to the study were held. There were several aims to these meetings:

- to flag up recruitment issues at an early stage;
- to iron out problems collaboratively with the midwives;
- to discuss any mental health issues; and
- to convey gratitude to CMWs for their support.

This worked well with some of the groups, but not all, and it became apparent that not all midwives were positively inclined towards the subject of the study. Concern about possible bias led to a review of individual midwives' recruitment figures. It became apparent that those midwives who seemed to find the study problematical did not recruit at all, thereby making it unlikely that selection bias was introduced.

Women were therefore either not approached at all or else were recruited by a CMW with no bias against the study.

3.2.2.7 Research governance and ethical approval

The study was approved by LWH's R&D committee and by the local REC; approval was also obtained from the local medical committee, and all GPs in Liverpool were informed of the possibility of one of their patients being approached about the study.

Because of child protection concerns involving disclosure of abuse, the study inclusion age was 18 and above. However, despite this, there was a possibility that child protection issues might arise, and these were discussed in supervision meetings. During the entire duration of the study, there was only one incident where there were child protection concerns: in the course of an interview a study participant disclosed a worry about her younger sister being in contact with a perpetrator; with her permission, the concerns were passed on to social services and the police.

3.2.2.8 Sample results and response rate

Figure 3.1 shows response rates at various stages of the study. Of 1,307 women screened for inclusion into the study, 16 (of whom 14 had consented to participation and 2 had refused) either had a loss of pregnancy during the study period or lost the baby and were therefore excluded from further analysis.

Respondents

- 1,291 women were eligible for inclusion and classed as the study sample.
- 1,029 women (79.7 per cent of study sample) consented to participate (consenters).
- 821 women returned the antenatal (time 1) questionnaire (64 per cent of study sample; 80 per cent of consenters).
- 695 women returned the 6-week postnatal (time 2) questionnaire (54 per cent of study sample; 68 per cent of consenters; 85 per cent of those responding at time 1).
- 506 women returned the 12-week postnatal (time 3) questionnaire.
- 88 women who had returned their antenatal (time 1) questionnaire had not been sent a 12-weeks postnatal (time 3) questionnaire. Hence, the response rate at 12 weeks postnatal cannot be calculated from the overall sample. Of the 733 women

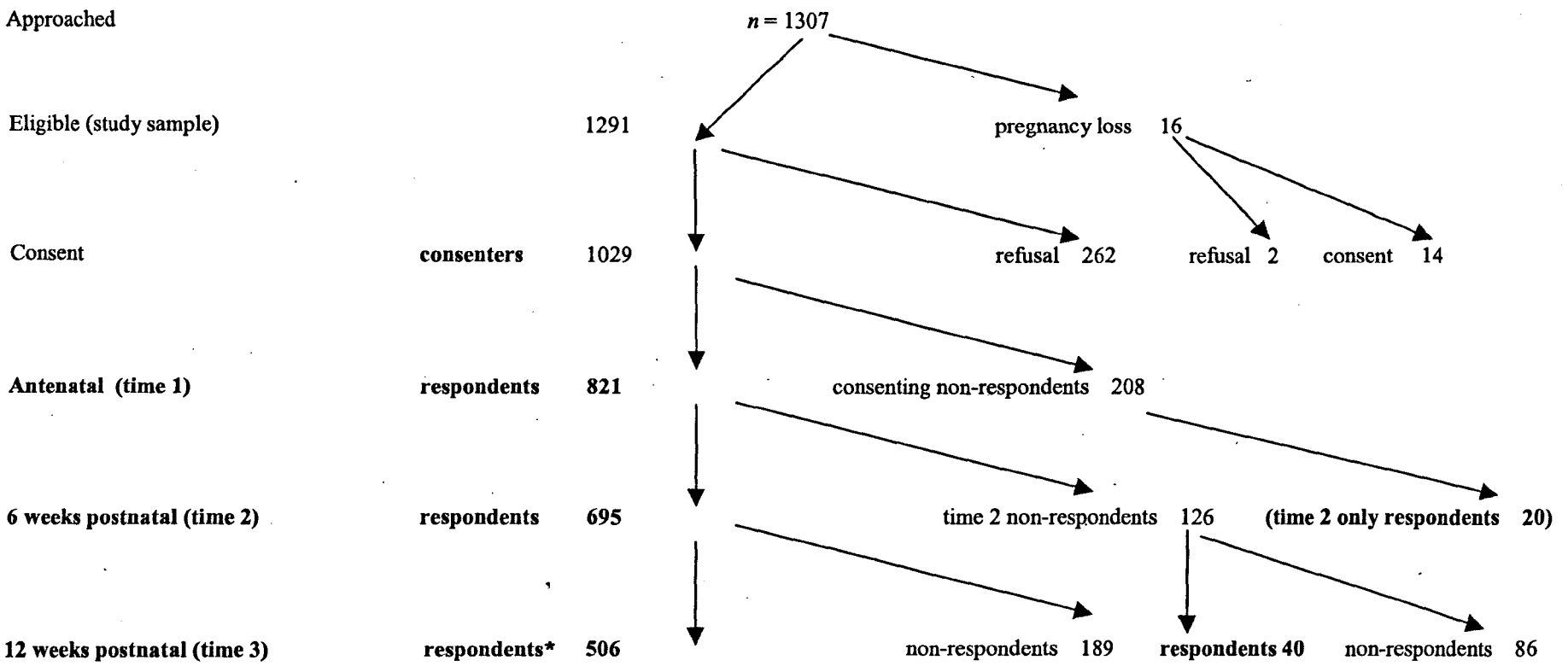
sent a 12-weeks postnatal (time 3) questionnaire, the response rate was 74 per cent.

- 622 women completed all sub-sections of the questionnaires relevant to this study before 39 weeks gestation and form the sample for time 2 results. Of those, 450 women completed the EPDS section of the 12-weeks postnatal (time 3) questionnaire, and therefore form the sample for time 3 results.
- 68 (87 per cent) of 78 women invited agreed to participate in diagnostic SCID interviews.

Refusers

Of the 1,291 eligible women (study sample) 262 women (20 per cent) declined participation and are classed as refusers. Of these, 184 provided some data, such as age, occupation and postcode, whereas 38 did not wish to provide any basic demographic data at recruitment stage. The consent status of forty women was not known and basic demographic data were not provided. Analyses of this data are included in the results on non-participants (see 4.2).

Figure 3.2.2.8 *Sample recruitment*



* *Note*: 88 women who had returned the antenatal (time 1) measure had not been sent the 12 weeks postnatal (time 3) measure

3.3 Measures

In what follows an overview is provided of the antenatal and postnatal measures used in this study, their components and the constructs they tested; table 3.3 gives a summary.

3.3.1 Antenatal (time 1) questionnaire in late pregnancy (>24 and <39 weeks gestation) (Appendix A)

3.3.1.1 Socio-economic, educational and health variables

Each participant was asked to answer questions about herself and her pregnancy, her partner, employment and educational attainment, previous physical and mental illnesses, current alcohol and cigarette consumption, socio-economic details and family composition during childhood.

Townsend score

Postcodes were obtained and, with the help of a public health database, converted to Townsend scores (Phillimore *et al.* 1994), a ward-based measure of material deprivation, combining four variables selected to reflect distinctive aspects of material well-being: unemployment, car-ownership, non-owner occupation and household overcrowding. This score is commonly used in public health and is derived by converting postcodes, thereby providing an approximation of social deprivation for participants in the study, as it applies to those living within the specified area of a ward rather than a particular household. It must be remembered that classifications based on areas rather than individuals can fall foul of the ecological fallacy – assuming unjustifiably that a relationship which appears at the small-area level reflects an underlying relationship at the level of the individual (Drever & Whitehead 1995).

Predictive Index (PI)

The study was designed to investigate also a relationship between a predictive tool of PND, the Predictive Index (Cooper *et al.* 1996), and childhood adversities. This,

however, is not the focus of this dissertation, so that no analyses regarding childhood variables and the PI are presented here. Some items from the PI were, however, used in the description of the sample and to obtain information on any history of previous depression (PrevD) prior to pregnancy.

The PI was validated on a large postnatal sample of nearly 2,000 women (Cooper *et al.* 1996). At a score above 27 the risk of PND was 35 per cent, and more than one-third of those who were to become depressed scored in that range. Its seventeen weighted variables group into those concerned with the emotional and physical experiences of pregnancy, previous experience of mood disorder, the quality of close relationships and socio-economic information.

3.3.1.2 Antenatal depression

Edinburgh Postnatal Depression Scale (EPDS)

This is a brief self-administered screening instrument consisting of 10 questions, each rated on a 4-point scale (0–3), standardised for postnatal (Cox *et al.* 1987) and antenatal use (Green & Murray 1994). Individual scores range from 0 to 30, the higher scores indicating higher levels of depression, and a recommended cut off of 12–13 can be used to create a dichotomous variable of presence or absence of depression (see 1.1.1.1).

Concerns have been raised that in its antenatal use, question 10 (assessment of suicide risk) may be misinterpreted to mean accidental injury (Green & Murray 1994). Because of these concerns, the wording of question 10 in the *antenatal* (time 1) questionnaire in this study was amended slightly, to read '*I have thought of harming myself*' instead of the original '*The thought of harming myself has occurred to me*'.

3.3.1.3 Relationship with parents

Parental Bonding Instrument (PBI)

Recalled parental care and control prior to the age of 16 were assessed using the PBI (Parker *et al.* 1979). This 25-item self-report instrument consists of 12 items for the measurement of care and 13 for control, each item to be rated on a 4-point scale (0–

3). Many questionnaire-based large-scale studies use the PBI because it is reported to be psychometrically robust (Parker 1989), stable over time, and unaffected by age, social class or gender (Parker 1989; Parker 1990), and it does not appear to be strongly affected by current mood (Parker 1981; Gotlib *et al.* 1988; see 2.2.2.3.2).

3.3.1.4 Abusive childhood experiences

CSA

The sexual abuse items were the same as those used in previous population-based studies (Mullen *et al.* 1993; Hill *et al.* 2000). Participants were asked about unwanted sexual experiences prior to the age of 16 by a person at least 5 years older. Questions differentiated:

- non-contact abuse;
- contact abuse involving touching of participant's or perpetrators' genitalia;
- contact abuse involving attempted or completed intercourse; and
- any other unspecified contact abuse.

Women were asked to provide details of the frequency of the abuse (never, once, several times, often), their relationship with the perpetrator, where the perpetrator lived, the perpetrator's age and gender, and the participant's age at time of the first instance of abuse.

Physical abuse

A review revealed that most extant measures of childhood physical abuse are not entirely satisfactory because they use subjective assessments (e.g. Fergusson *et al.* 2000), for which reason they are not commonly used, as outlined in chapter 2 (2.2.2.2). A comparison of existing measures suggested certain kinds of behaviour that need to be covered (Drossman *et al.* 1990; Parker *et al.* 1997). Because it has been established that the accuracy of the results of questionnaires which assess abuse retrospectively increases with the number of questions and their level of detail (Fergusson & Mullen 1999), a scale was devised which appeared to satisfy those concerns. In particular, the aim was to devise a tool that established *objective* behaviours by providing descriptions of those behaviours that could be rated according to the frequency with which they occurred. Additionally, it was important

to ascertain the presence of physical abuse in a fashion to similar the other measures of childhood adversity, such as parental behaviour and CSA, by using the same age range.

In this study, therefore, physical abuse prior to the age of 16 was assessed by enquiring about the frequency (never, once, several times, often) and severity of instances of physical abuse perpetrated by a responsible adult, such as blows to the face using the hand or blows with an object; the exertion of force sufficient to cause injury; injuries requiring medical attention. The age at which physical abuse first occurred, the relationship with the perpetrator and the gender of the perpetrator were recorded.

3.3.1.5 Satisfaction with questionnaire and free text

Women had the opportunity to provide further thoughts about the relationship between childhood experiences and their current pregnancy and to provide feedback on the questionnaire.

3.3.2 Postnatal (time 2) questionnaire at six weeks postpartum (Appendix B)

3.3.2.1 Delivery outcome

Questions covered details about the baby (age, gestation at delivery, weight), delivery mode and complications, and time spent by mother and baby in hospital.

3.3.2.2 PND

EPDS

This scale was chosen to assess presence of depressive symptoms (see 3.3.1.2 and chapter 1). The original wording was retained for question 10 in the *postnatal* version.

3.3.3 Postnatal (time 3) questionnaire at 12 weeks postpartum (Appendix B)

This was the same as the six weeks postnatal (time 2) questionnaire, including the EPDS as the outcome measure of PND, with additional opportunity to provide thoughts about participation in the study, satisfaction with the questionnaire and health care received.

Table 3.3 *Overview of instruments used in study*

Instrument	Acronym	Time when used	Construct
Parental Bonding Instrument	PBI	Antenatal	Reported perceived parental care and control
Physical abuse questions	PA	Antenatal	Reported childhood physical abuse experiences
Childhood sexual abuse questions	CSA	Antenatal	Reported childhood sexual abuse experiences
Townsend Score	Townsend	Antenatal	Current socio-economic status
EPDS	EPDS	Antenatal; 6 weeks pp; 12 weeks pp	Depressive symptoms

3.4 Data management

3.4.1 General data management

All women registered by CMWs were allocated a personal study number and their details entered into a specifically designed (MS Access) database. Missing or erroneous information about participants was corrected by checking the hospital database or contacting patients' GPs. The database was programmed to automatically generate lists of women who were due to be sent questionnaires at the relevant time points. Data from returned questionnaires were entered into the Access database. Data entry was checked for correctness and later imported into SPSS PC (SPSS inc. 2001) for analysis.

3.4.2 Missing data

Rules were devised for the handling of missing data. For EPDS and PBI, whose items are scored on a straightforward Likert scale of 0–3, missing data was pro-rated by calculating an item mean for all the items provided, which was then divided by the number of items and added to the total. For any participant, a maximum of 3 missing items were pro-rated for the EPDS, and a maximum of 5 for the PBI. The rules for the predictive index were slightly more complicated due to each item being individually weighted. Therefore, the total score was divided by the total possible score (63) minus the maximum score of the missing item. This proportion was then multiplied by the maximum score for the missing item. The result was added to the individual's total sum on the predictive index. Up to five missing items were allowed to be pro-rated on the PI. Table 3.4.2.1 provides an overview of the number of cases with missing data needing to be pro-rated.

Table 3.4.2.1 *Overview of number of measures with cases of missing data pro-rated*

Measure		Number of cases pro-rated (%)
EPDS		
	antenatal	14 / 821 (1.7%)
	6 weeks postnatal	7 / 695 (1.0%)
	12 weeks postnatal	5 / 506 (1.0%)
PBI		
	Maternal care	35 / 821 (4.3%)
	Maternal control	27 / 821 (3.3%)
	Paternal care	32 / 821 (3.9%)
	Paternal control	32 / 821 (3.9%)
PI		110 / 821 (13.4%); (of which 109 cases had 1 item missing, and 1 case had 2 items missing)

3.4.3 Comparison data

A retrospective anonymised data set obtained from LWH provided for all women who had booked into antenatal clinic during the period of the study information on booking date, postcode, age, gestation at delivery and GP. This data set was prepared so that the geographical area would correspond to that of the study sample (i.e. women were deleted if their postcode or GP was outside of Liverpool) and Townsend scores were obtained. This data set was later merged with the study sample for comparison of the two with respect to bias (see 4.1).

3.5 Variables

3.5.1 Antenatal and postnatal depression

Postnatal depression at 6 and 12 weeks postpartum was the dependent variable. The forms of depression investigated in this study were:

- previous depression (PrevD) (women who had reported no depression prior to this pregnancy)
- antenatal depression;
- postnatal depression at 6 weeks postpartum (time 2);
- postnatal depression at 12 weeks postpartum (time 3);
- persistent PND (women who had depression at both 6 and 12 weeks postpartum);
- transient PND (women who had depression at 6 weeks postpartum only);
- newly emerging PND (women who had reported neither previous depression [PrevD] nor AND).

AND and PND scores were not normally distributed; therefore, depression was treated as a binary variable (present/absent).

3.5.1.1 Diagnostic interviews for depression

In order to subsequently establish the best cut-off for the EPDS in this study sample, 68 diagnostic postnatal SCID interviews (First *et al.* 1994) on a random sub-sample of the study population were conducted by 2 interviewers who were blind to postnatal (time 2) EPDS scores. Of the 78 women approached postnatally for participation in interviews, 10 refused (12.8 per cent). (For the selection of this sub-sample, see 3.2.2.4). Interviewers were instructed to ascertain:

- a diagnosis of depression at the time of completion of the EPDS; and
- a diagnosis of depression at any time following delivery.

Tables 3.5.1.1.1 and 3.5.1.1.2 show the statistical correlations between diagnoses derived by interview and different EPDS cut-off scores, the higher scores indicating increased levels of depressive symptoms. When comparing SCID diagnoses with EPDS scores, one needs to bear in mind that the SCID uses a two-week time frame for symptoms to be required for a diagnosis of depression, whereas the EPDS requires the presence of symptoms only for the past week. Therefore, results are shown for a diagnosis of SCID depression made at the time of the completion of the

EPDS (table 3.5.1.1.1) and of SCID depression covering any time of the postnatal period (table 3.5.1.1.2). The mean baby age at interview was 9.92 weeks (range 7–15; *SD* 1.73), compared to the mean baby age at completion of the EPDS of 6.93 weeks (range 6–10; *SD* 1.23). When the time scale for SCID depression was identical with the time when the EPDS had been completed, a cut-off score of 12–13 on the EPDS yielded the highest kappa of 0.55, indicating moderate agreement (Landis & Koch 1977) between the two measures. This finding suggested that this widely used cut-off (Cox *et al.* 1987; Murray & Carothers 1990) was appropriate for this study and it was therefore chosen for analyses of antenatal and postnatal EPDS scores. Observer-rated agreement was stronger when the time frame for SCID depression was extended to include any time within the postnatal period up to the interview (table 3.5.1.1.2), with a sensitivity (proportion of positives correctly identified) of 68 per cent (17 of 25) and a specificity (proportion of negatives correctly identified) of 93 per cent (40 of 43). As shown previously (see 1.1.1.1), this compares well with the findings of 67.7 per cent and 95.7 per cent, respectively (Murray & Carothers 1990).

Table 3.5.1.1.1 *Agreement between DSM-IV diagnosis of depression (SCID) at time of completion of EPDS and different EPDS cut offs*

EPDS cut-off	**SCID diagnosis at time of completion of EPDS (n)**		**Kappa**	**Asymptomatic standard error**	**Approximate T**	**Approximate significance *p***
	Absent	**Present**				
11–12			0.45	0.100	4.26	<0.001
below	38	16				
above	1	13				
12–13			0.55	0.104	4.88	<0.001
below	42	12				
above	1	13				
13–14			0.53	0.116	4.53	<0.001
below	45	9				
above	3	11				

Kappa, asymptomatic standard error and approximate T derived from 2x2 contingency tables (total $n = 68$)

Table 3.5.1.1.2 *Agreement between DSM-IV diagnosis of depression (SCID) at any time during postnatal period and different EPDS cut offs*

EPDS cut off	SCID diagnosis at any time during postnatal period up to interview		Kappa	Asymptomatic standard error	Approximate T	Approximate significance *p*
	Absent	Present				
11–12			0.59	0.097	5.09	<0.001
below	37	11				
above	2	18				
12–13			0.64	0.098	5.32	<0.001
below	40	8				
above	3	17				
13–14			0.65	0.102	5.32	<0.001
below	43	5				
above	5	15				

Note: Kappa, asymptomatic standard error and approximate T derived from 2x2 contingency tables (total $n = 68$)

3.5.2 Parental care and control

Scores of the PBI variables for maternal and paternal care and control were scewed (see 4.3.1 and table 4.3.1.3). Binary variables were therefore created in line with Rodgers (1996) and Hill *et al.* (2000) to establish low care and high control variables. Thus for low care a threshold was chosen that identified the lowest quartile for care (25/26 for maternal care and 21/22 for paternal care) and for high control the threshold was chosen that identified the highest quartile for control (12/13 for maternal control and 14/15 for paternal control).[1]

3.5.3 CSA

CSA was treated as a binary variable and rated as present if a participant indicated any presence of contact sexual abuse prior to the age of 16 perpetrated by a person older by at least 5 years.

1 The cut-offs for quartiles were:
Maternal care: quartile 1 (lowest): 0–25.9; quartile 2: 26–32.9; quartile 3: 33–35.9; quartile 4: 36+
Maternal control: quartile 1: 0–4.9; quartile 2: 5–8.9; quartile 3: 9–12.9; quartile 4 (highest): 13+
Paternal care: quartile 1 (lowest): 0–21.9; quartile 2: 22–29.9; quartile 3: 30–33.9; quartile 4: 34+
Paternal control: quartile 1: 0–5.9; quartile 2: 6–9.9; quart 3: 30–33.9; quart 4 (highest): 15+)

3.5.4 Physical abuse

Physical abuse was also treated as a binary variable and coded as present if participants reported that it had occurred one or more times prior to the age of 16.

3.5.5 Current social deprivation

Because the Townsend distribution was scewed, with a range from –4 to +13 (mean = 4.0; median = 5.0; mode = 0; *SD* 3.97) and, in line with MacArthur *et al.* (2002), the Townsend score was used as a categorical variable. However, as quartiles did not provide even numbers in each section, analyses were instead based on quintiles,[2] quintile 1 and 5 indicating, respectively, the least and the most deprived. As analyses of the persistence of PND were performed on a sub-sample of the overall sample (see 4.5 and 4.6), using Townsend quintiles would not have provided adequate numbers in each cell. Therefore, for analyses with low sample size, a binary variable for Townsend was created, consisting of the top two quintiles (most deprived) and lowest three (least deprived).

3.5.6 Teenage pregnancy

A variable for teenage pregnancy was defined as those women who were below the age of 20 at the completion of the antenatal (time 1) questionnaire.

3.5.7 Continuous variables

Age, gestation, Townsend score (indicating the level of social deprivation) and EPDS score were used as continuous variables to describe the properties of the sample.

---

2 Ranges for the quintiles are as follows:
Quintile 1 (least deprived): range –4 to –1;
Quintile 2: range 0 to +3;
Quintile 3: range +4 to +5;
Quintile 4: range +6 to +7;
Quintile 5 (most deprived): range +8 to +13.