Should women be screened for postnatal depression? Exploring the effects of undiagnosed maternal mental health problems on child development.

Kerry Jane Bell

Doctor of Philosophy

University of York

Health Sciences

2014

Abstract

Background: Evidence of a relationship between maternal postnatal depression and child development is consistently growing, but there has been no distinction between depression that is clearly diagnosed and treated, and depression that is never identified by health professionals. Previous models assessing the cost-effectiveness of screening for postnatal depression have been unable to account for child outcomes and the effects of undiagnosed maternal depression due to a lack of research in this area. Without these outcomes, screening for maternal postnatal depression is not currently considered to be cost-effective.

Methods: Longitudinal survey data from the Millennium Cohort Study is used to explore the differential effects of undiagnosed and diagnosed maternal depression on child cognitive and behavioural development over time, and to re-examine whether screening for postnatal depression could be considered cost-effective once longer term child outcomes are included.

Results: Depression that is undiagnosed has a substantial effect on the behavioural development of children. Children of mothers who are depressed but not diagnosed at 9 months are at least equally likely as those of mothers with diagnosed and treated depression to have behavioural problems later on in childhood. Identifying and treating maternal depression showed some short-term beneficial effect for child behavioural development up to age 5, but this was not maintained at age 7. Higher levels of persistent depression were identified in women who were diagnosed and treated for depression and this persistency was found to have an additive effect on child outcomes, with longer-term maternal mental health problems much more strongly associated with child outcomes than postnatal depression alone.

Conclusions: This research highlights the limited success of current treatments for maternal depression, both in benefiting child development and providing long-term symptom remediation for mothers. As current treatments lack benefit for children over the longer-term, the recommendation that screening for postnatal depression appears not to be cost-effective remains unchanged.

Table of Contents

List of Figures

List of Tables

Chapter 6

Chapter 7

List of Appendices

Acknowledgements

There are a number of people I would like to thank who have supported me in carrying out my research. Firstly I would like to thank my supervisor, Professor Karen Bloor, who has been a constant source of support and encouragement throughout my PhD, providing valuable feedback and discussion with much fun and laughter along the way. Secondly I would like to thank Dr Catherine Hewitt for imparting her statistical wisdom and advice upon me, and for always taking the time to explain the numerous statistical methods I have had to learn in order to conduct this research. Finally I would like to thank Professor Alan Maynard and Professor Kathleen Kiernan for their much appreciated contributions to discussions surrounding my work and for providing feedback and expertise along the way as members of my thesis advisory panel.

A special thanks is due to Dr Jelena Mirkovic for encouraging me to believe in myself and giving me the confidence to apply for a PhD studentship. Without her support and encouragement I would not be where I am today.

I am grateful to Professor David Torgerson and my other colleagues in York Trials Unit who have all been very supportive of me whilst I have been completing my PhD over the past few months.

I am grateful to all of my friends, old and new, for all of their support over the last few years and for encouraging me to keep going when things got tough. I would like to give particular thanks to Kathryn Williams, my oldest and dearest friend and a source of unwavering love and support, as well as Emily Bailey, Lily Lewis, Holly Essex, Laura Jefferson and Noortje Uphoff.

Finally I would like to thank my family, especially my parents Colin and Michelle and my brother Reece for their support and encouragement, and my partner Alex for his love and reassurance over the last few years.

Author's Declaration

I confirm that the research presented within this thesis is my own work. It has not been submitted for an award elsewhere. To the best of my knowledge, this thesis does not contain material that has been published or written elsewhere by another person, except where acknowledgement and appropriate reference is made in the text.

Chapter 1: Background and focus of the thesis

1.1 Introduction

The overarching context of this thesis is the importance of the early years for child development over time and the costs and benefits associated with early intervention. Specific focus will be placed on how exposure to maternal depression during the early stages of a child's life affects developmental outcomes. The case for early intervention for maternal depression is addressed and the potential costs and benefits are explored. This chapter provides a background to the thesis, describes the overall objectives, and outlines the structure of the thesis.

1.2 Background: The importance of the early years

Health and wellbeing during childhood shape future health and development. Determining why some individuals do better than others in both the early years and later life is a key issue for policy makers and is crucial in attempts to reduce future inequalities (Cunha and Heckman, 2007, Pugh and Duffy, 2013). Early childhood provides the foundations for later development with educational and social experiences during this period being important for developmental outcomes through to adulthood (Schaffer, 2000). It is well established that learning starts during early life, long before formal education begins, and continues throughout the life course (Heckman, 2007). This line of thinking has informed many early childhood based interventions such as Sure Start in the UK (Glass, 1999) and its predecessor Head Start in the US (Garces et al., 2002) which are premised on the idea that cognitive stimulation and supportive environments are needed to support a series of developmental steps in early infancy, with the view of preventing later cognitive delays and inequalities. Head start began as a short-term summer program in 1965 with a predominantly African-American sample but was expanded to serve both African-American and white children by 1966. Head Start is comprised of a number of individual programs, all of which aim to address different issues in child development. For instance, Early Head Start promotes healthy prenatal outcomes, healthy families and infant and toddler development beginning as

early as birth. Head Start programs in later childhood promote child health and family-community partnerships to offer parents opportunities and support to identify and meet their own goals in terms of education and employment. Evaluation studies of the Head Start program has identified short-term benefits, as indicated by test scores (Barnett, 1995) and more lasting improvements in non-cognitive skills such as persistence and dependability that are important for future success in life (Heckman et al., 2000).

Founded on this principle, Sure Start was announced in the UK in 1998 and aimed to improve health and learning outcomes in children as well as social and emotional development and parenting skills. Sure Start Children's Centres provide a number of services such as; integrated early learning and childcare, drop-in activity sessions for children, family support, including support and advice on parenting, information about services available in the area such as antenatal and postnatal support, information and guidance on breastfeeding, health and nutrition, smoking cessation support, and speech and language therapy and other specialist support, links with jobcentres to encourage and support parents and carers who wish to consider training and employment (Department for Education and Skills, 2007). Evaluative research has shown Sure Start to have positive effects on parenting (Belsky et al., 2006), child behavioural outcomes and child health outcomes (National Evaluation of Sure Start team, 2012).

Almost every aspect of the developmental process is affected by the environments and experiences that are encountered in the early years of life (Shonkoff and Phillips, 2000). Consequently, environments that do not facilitate adequate stimulation and support for the child from a young age place children at an early developmental disadvantage, with ability gaps between advantaged and disadvantaged children being shown to open up as early as age four (Carneiro, 2003).

Development can be influenced by a range of factors and circumstances. One of the most consistently recognised influencers of development is family economic hardship and poverty, particularly with regards to the dichotomy between being raised in poverty versus not being raised in poverty (Shonkoff and Phillips, 2000, Kiernan and Mensah, 2009). Research using British birth cohort studies has shown that even by the age of 2 years, children from lower socio-economic backgrounds have lower cognitive scores which can impact on later educational attainment (Feinstein, 2003). Other research has shown that economically disadvantaged mothers are more likely than the more

advantaged to experience psychological problems, particularly depression (Reading and Reynolds, 2001) which can in turn have detrimental effects on children (Grace et al., 2003, Goodman et al., 2011). One study by Kiernan and Mensah (2009), attempts to distinguish the differential effects of maternal depression and poverty on child outcomes. A sequential series of logistic models were derived, firstly assessing each of the variables (poverty and maternal depression) independently, and then by adjusting for the each other and finally by adjusting for each other alongside a set of known covariates (e.g. family status, child gender, maternal age, maternal qualification). A reduction in effect sizes was identified for each additional adjustment of the model. In the fully adjusted models poverty was significantly associated with both child learning delays and behaviour problems, whilst maternal depression was only significantly associated with child behaviour problems.

Often relating to poverty, family structure, in terms of single or dual parenting, is highly important for understanding how household resources affect child development. National survey research, as well as smaller scale studies both in the US and the UK, have consistently reported that children in single parent families have a higher probability of problems in health, educational achievement and behavioural development problems than children who have two married parents (Chaimay et al., 2006, Goodman et al., 2011). Single parenting often reduces household income, places mothers at increased risk of depression (Crosier et al., 2007, McKay, 2002), and decreases parents' ability to supervise their children and participate in their activities (Levine-Coley, 1998).

Parental employment also plays an important role in determining the resources available to a family, and the amount of time parents are able to spend with their children (Shonkoff and Phillips, 2000). Early return of mothers to full time employment post childbirth has been shown to be a risk factor for depression, particularly amongst single parents and may have effects on child cognitive development (Brooks-Gunn et al., 2002, Han et al., 2001, Waldfogel et al., 2002). Since the 1980s, employment of mothers has been increasing (Fagan and Norman, 2012), with the employment rate for mothers with a child aged under the age of five years reaching 58 per cent in 2008 (OECD, 2011).

1.3 The family environment

Although there are many factors that can affect child development, families, particularly parents, play a powerful role in shaping adult outcomes, particularly cognitive and socio-emotional abilities (Heckman, 2006). Almost all children form close emotional bonds, or attachments, to care givers during the early years of life, especially, in the majority of cases, the mother (Shonkoff and Phillips, 2000). Attachments, as is discussed in Chapter 2, are thought to result from close contact between infants and caregivers who can protect them and guide their development, as well as creating a foundation for future interactions and other behaviours throughout the life course (Grych & Fincham, 2001, Dogra et al., 2002).

In a review of 22 studies (Ranson et al., 2008) positive parent-child relationships were found to affect socio-emotional competence, to be important for developing other relationships, such as with peers, and to facilitate the development of personality traits such as social competency, sociability, friendliness, and co-operation skills. In contrast, insecure attachments were shown to be related to aggressive and negative affect behaviours by the age of four years. Links were also identified between attachment and child cognitive outcomes lasting into adulthood. Although this review did not use rigorous scientific methodology through a systematic search strategy to identify literature, and thus could be missing valuable research presenting opposing findings, the included studies were assessed for quality and thus present a strong case for the importance of early relationships.

More recent literature further supports this notion, describing how the parent-child relationship affects educational and behavioural outcomes (Ali, 2011), and how parenting relates to the development of comorbid problems in children with ADHD (Deault, 2009). Quality parenting is thus highly important for child outcomes.

1.4 Postnatal depression; classification and clinical diagnosis

Clinical diagnosis of psychological conditions is generally based on one of two diagnostic manuals; The Diagnostic and Statistical Manual of Mental Disorders (DSM-V) (American Psychiatric Association, 2013) or the International Classification of Diseases 10 (ICD-10) (World Health Organisation, 1992). Neither currently recognises

postnatal depression as an individual condition, instead, patients must meet the criteria for a major depressive episode with an onset in pregnancy or within four (DSM-V) to six weeks of delivery (ICD-10). The DSM-V states that depression is present when an individual experiences five or more of the following nine symptoms on a daily basis over a two week period: depressed mood, loss of interest or pleasure in usual activities, change in weight or appetite, insomnia or hypersomnia, psychomotor retardation or agitation, loss of energy or fatigue, feelings of worthlessness or guilt, impaired concentration or indecisiveness or final recurrent thoughts of death or suicidal ideation or attempt. The ICD-10 follows very similar criteria.

However, organisations specialising in serving postnatal mothers have emphasised that 4-6 weeks may be too short and the onset of postnatal depression may occur up to one year after birth (PANDAS UK, 2014). The Scottish Intercollegiate Network (SIGN) also specify in their guidelines that depression occurring during the first postnatal year should be considered postnatal depression (SIGN, 2012) and self-help guidance from BUPA UK further reinforces this (BUPA, 2011).

Previous research in this field has also highlighted the importance of maternal depression at time points in the first postnatal year beyond the first four to six weeks after birth (Kiernan and Mensah, 2009, Kiernan and Huerta, 2008, Hobcraft and Kiernan, 2010), and leading theories in child attachment discussed in Chapter 2, place emphasis on the whole first postnatal year and beyond (Bowlby, 1969, Ainsworth, 1979).

In addition to these guidance documents for clinical diagnosis, there are also a range of self-report questionnaires which can be used to inform case detection. These are described in Chapter 3, section 3.2.2.)

1.5 Cause of maternal postnatal depression

Depression as a whole has no established mechanism that can describe accurate aetiology for all individuals (Belmaker and Agram, 2008). The primary competing arguments centre on biology and psychosocial/ psychological factors. Postnatal depression is also dominated by these same arguments. There is a considerable amount

of literature on the risk factors for postnatal depression attracting several meta-analyses and systematic reviews (O'Hara and Swain, 1996, Robertson et al., 2004). These have highlighted several key risk factors for the development of postnatal depression in line with the biological and psychological approaches. The biological approach centres on hormonal changes, suggesting that changes in oestrogen and progesterone levels following delivery may affect mood. Randomised controlled trials of hormonal therapy as a treatment for depression reported some success thus suggesting hormone changes may contribute to depression in some women (Gregoire et al., 1996) but this is not universally accepted (Harris et al., 1996).

The psychological approach offers two primary explanations, psychiatric risk factors and stress (Milgrom et al., 2008). Both family and personal history of depression have consistently been reported as risk factors (O'Hara and Swain, 1996). Additionally, depression onset during pregnancy is also thought to contribute to continued depression or recurrence (if previously resolved) during the postnatal period (Heron et al., 2004). More generally, personality traits such as high interpersonal sensitivity and low self-esteem may also contribute, possibly relating to self-doubt in aspects of motherhood (Ritter et al., 2000).

Moreover than psychological risk factors, stress is thought to play a pivotal role in the development of both general depression and postnatal depression. Recent stressful events in the year prior to pregnancy can predict depression in pregnancy and the postnatal period (Rubertsson et al., 2005). Additionally, the perception of a lack of support in caring for the infant may also cause stress (Honey et al., 2003), in addition to more general concerns such as income and housing (Patel et al., 2002). In qualitative work conducted by McIntosh (1993), new mothers who had experienced postnatal depression were asked their opinions of the cause. The most commonly cited 'causes' of depression were aspects of motherhood (74%), with women citing boredom, loneliness and difficulties in coping with a new baby as particularly stressful, and financial concerns (71%). Only 11% of women thought their symptoms were likely due to biological factors suggesting that postnatal depression is generally perceived as a psychological problem by women. If these factors believed by women to be causal are in fact the basis of their depression, this presents issues for applying treatment as will be discussed later in this thesis.

1.6 Prevalence of maternal postnatal depression

Despite being published over a decade ago, the most predominant estimate for the prevalence for postnatal depression in women is that generated through a large scale meta-analysis of 59 studies conducted by O'Hara and Swain (1996), which placed prevalence at 13% of the population. However, maternal depression may be an underdiagnosed condition (Hearn et al., 1998, 4Children, 2011), thus the actual rates could be much higher. Results from a cross-sectional study comparing the opportunistic diagnostic abilities of health professionals (general practitioners, health visitors and midwives) to results of a self-reported Edinburgh Postnatal Depression Scale (EPDS) questionnaire showed a significant failure of health professionals at detecting symptoms of maternal depression during routine postnatal follow up care (Hearn et al., 1998), with the primary health team collectively missed 57% of women with high levels of self-reported symptoms. This suggests that relying on opportunistic identification of depression by health care professionals alone is insufficient and likely to result in a high number of undiagnosed cases, highlighting the value of self-report measures as useful screening tools. More recently a report published by the charity '4Children' in 2011, which surveyed 2318 mothers on their experiences with postnatal depression, documented that 33% of new mothers with more than one child and 26 % of first time mothers reported suffering from postnatal depression. The report further identified that despite a large proportion of women acknowledging having symptoms, only 54% of multiparous women (those with more than one child) and 42% of primiparous women (those with only one child) actually sought professional help, indicating that a large proportion of depressed women are not reporting their symptoms to a doctor. This confirms that the actual prevalence of postnatal depression is likely to be considerably higher than previous estimates.

1.7 Importance of maternal mental health

Maternal mental health may affect the quality or style of parenting and is increasingly being recognised as important for shaping child development. Impaired maternal mental health as a consequence of postnatal depression has been shown to have lasting deleterious effects on child development including poorer cognitive outcomes (Hay et al., 2001) and increased behavioural disturbances (Kiernan and Mensah, 2009).

There has been a recent increase in the interest in maternal postnatal depression, attracting media attention through articles and TV documentaries (BBC documentary: Depression, teen mums and me, March, 2013).

Systematic and non-systematic literature reviews focusing on child development have highlighted the importance of maternal depression. Breinholst et al. (2012) emphasise the importance of maternal mental health, finding a strong association between maternal mental health issues and child mental health outcomes, such as anxiety onset. It was also noted that the presence of impaired maternal mental health impeded the treatment of anxiety in children. DeWalt and Hink (2009) reported that children of mothers with more depressive symptoms had more depressive/ withdrawn symptoms themselves, but only in the presence of low maternal literacy indicating that although maternal mental health is important, other variables can mediate the effects it has on children's outcomes.

Two systematic reviews have collated the evidence on the relationship between maternal mental health and child outcomes (Goodman et al., 2011, 2007, Grace et al., 2003). Grace et al. (2003) reviewed 13 papers exploring the effects of maternal postnatal depression on child cognitive and behavioural development. Small effects were found on particular aspects of children's cognitive development such as language and IQ. These were mostly found on a short-term basis with longer term studies of up to 5 years showing no significant differences. Although small, questionable differences were found for cognition, long-term effects extending up to 5 years were found for behavioural outcomes with children of depressed mothers performing consistently and significantly worse than children of control mothers. It was further suggested that mothers with more persistent depression had children with greater behavioural problems implying increased effects with cumulative exposure to symptoms of maternal depression.

In contrast to reviewing studies based on only postnatal depression, Goodman and colleagues examined the effect of maternal depression at any time point on child outcomes conducting a large scale meta-analysis of 193 studies to examine the strength of the association between mothers' depression and children's behavioural problems or

emotional functioning. The effects found mirrored those identified by Grace and colleagues exploring just postnatal depression. Again, maternal depression was significantly related to higher levels of behaviour problems in children. Although these differences were significant, they were all relatively small in size and were significantly moderated by various other contexts.

Taken together these three studies depict the demonstrable effects of maternal mental health on child outcomes.

To update these findings and ensure currency in the topic, the search strategy employed in the most recent review on maternal health and child outcomes (Goodman et al., 2011) was used to identify current literature, published since 2009 when the original search was completed. The results are presented below.

1.7.1 Review findings

Sixteen recent studies were identified through the search which met the original Goodman et al. inclusion criteria but focused only on maternal postnatal depression rather than depression at other time points. Seven of these studies examined the effects of postnatal depression on child cognitive development and twelve examined child behavioural development. The age of assessment varied between studies with covering child outcomes from very early childhood at age 1, to late adolescence at age 17 years.

1.7.1.1 Cognitive development

The evidence for an association between maternal postnatal depression and child cognitive development was inconsistent. Whilst 3 studies reported little or no effects on measures of cognitive ability, assessed between ages 1.5- 5 years (Piteo et al., 2012, Kiernan and Mensah, 2009 a, b), the remaining 4 studies did note significant differences on measures of school adjustment and readiness (Kersten-Alvarez et al., 2012, Kiernan and Huerta, 2008) as well as language development (Kersten-Alvarez et al., 2012, Quevedo et al., 2011) during early childhood (ages 1-5 years) and GCSE performance in boys at age 16 years (Murray et al., 2010). All seven studies were comparable in terms of quality with each study analysing information from large samples of families

and engaging a good range of potential confounders in analytical methods. Given the lack of clarity on the association between maternal postnatal depression and child cognitive development it is apparent that more research is required to fully understand this relationship.

1.7.1.2 Behavioural development

In contrast to cognitive outcomes, the literature documenting the association between maternal postnatal depression and child behavioural development appeared far more consistent. The field of child developmental psychology distinguishes between "externalising" and "internalising" behaviour types and their associated disorders (Achenbach, 1978). Externalising behaviour problems are manifested in children's outward behaviour, particularly how they interact with the external environment, and can include disruptive, hyperactive, and aggressive behaviours (Campbell, Shaw, & Gilliom, 2000; Eisenberg et al., 2001). In contrast, internalising behaviour problems are those that are not necessarily visible externally but affect the child on a psychological level, these can include withdrawal, anxiety, inhibition, and depression (Campbell et al., 2000, Eisenberg et al., 2001, Hinshaw, 1987). Both types of behaviour problem were assessed within the literature and both were found to be associated with maternal depression.

Of the 12 studies exploring behavioural development as an outcome measure 11 reported a significant association, with behavioural development being affected throughout the whole of childhood from age 1 to age 17 across both genders (see Table A1.1, Appendix 1). Where studies examined externalising and internalising behaviours separately, a stronger association was generally noted between maternal depression and internalising behaviour problems, with 1 paper reporting this to be particularly evident in girls (Agnafors et al., 2013). The one isolated paper that did not report a significant association was identified as been low quality analysing only a small sample of non-diverse families.

These findings present clear and consistent evidence for the link between maternal postnatal depression and child behavioural development, identifying behaviour as an important outcome for current mother-child focused research.

1.7.1.3 Summary

These studies confirm the current importance of maternal mental health for child outcomes, particularly for behavioural development. Given the disparity in the literature surrounding the association between maternal postnatal depression and child cognitive outcomes, further research is merited, thus both aspects of child development (behaviour and cognitive) will initially serve as outcome measures for this thesis.

1.8 Early intervention

Early intervention can be the key to improving the long-term outcomes of disadvantaged children (Heckman, 2006). The longer the interval between onset of disadvantage and intervention, and the older the child, the more costly it is to remedy the disadvantage later (Heckman, 2008, Carneiro and Heckman, 2003). Early investment can result in greater long-term benefits producing a return on the intervention investment (Heckman, 2006). Over the life course, return from early intervention may be observed through increased educational attainment, often resulting in increased earnings, and improved health inclusive of physical and mental well-being. Returns can also be observed from a societal perspective through reduced crime and increased tax revenues gained via better income. Longitudinal studies find that personal benefits, social benefits and government savings associated with intervening early in a child's life often outweigh the costs (Karoly et al., 2005).

There is conflicting evidence considering the appropriateness of preventative strategies for postnatal depression, with a large scale Cochrane systematic review showing only marginal benefits of prevention based interventions (Dennis and Creedy, 2007). This systematic review identified 15 randomised controlled trials assessing both psychosocial and psychological interventions such as ante natal and post natal classes, professional and lay home visits, continuity of care and early post-partum follow up, and interpersonal psychotherapy. The various interventions were delivered by a range of health professionals such as nurses, community health visitors and doctors and the main outcome measure for all of them was onset of depression symptoms up to 24 weeks after birth. In 12 out of the 15 included trials post natal depression was defined as a score above an established cut-off point of greater than 12 points on the Edinburgh Post

natal Depression Scale (EPDS). The remaining trials employed other valid questionnaire measures and semi-structured diagnostic interviews. The findings from the systematic review noted that no significant prevention methodology was identified, although intense professional post-partum support showed some short-term benefits but these unfortunately were not maintained beyond 16 weeks which is when the frequency of visits was greatly reduced from weekly to monthly visits. Conversely to these findings, the review was later updated in 2013 leading to a complete change in perspective and the proposal that there may be some developing methods of preventing the onset of post natal depression which are effective. The update of the review was conducted by Dennis and Dowswell (2013) and included 28 trials comprising 17,000 women. Meta-analysed data from 20 of these trials (14, 727 women) showed that women who received a psychosocial or psychological preventative intervention were significantly less likely to develop post natal depression compared with those receiving just usual care (average RR 0.78, 95% confidence interval (CI) 0.66 to 0.93). The most effective interventions identified included intensive, individualised postpartum home visits provided by public health nurses or midwives (RR 0.56, 95% CI 0.43 to 0.73; two trials, 1262 women), lay (peer)-based telephone support (RR 0.54, 95% CI 0.38 to 0.77; one trial, 612 women), and interpersonal psychotherapy (standardised mean difference -0.27, 95%CI -0.52 to -0.01; five trials, 366 women). Professional and lay based interventions proved the most effective preventative interventions in reducing the risk of developing depression symptoms. Interventions commencing soon after birth also significantly reduced the risk of developing depression symptoms based on 12 of the included trials, equating to 12, 786 women (RR 0.73, 95% CI 0.59 to 0.90). Identifying mothers who were considered to be at increased risk of developing depression due to known risk factors such as poverty or being a single parent assisted in the prevention of postpartum depression (RR 0.66, 95% CI 0.50 to 0.88; eight trials, 1853 women). The methodological quality of the included trials was investigated and was reported to be of good to excellent though some trials did suffer from attrition. Six trials had losses to follow-up greater than 20% greatly increasing the potential risk of bias in these trials. The authors appropriately investigated the importance of these 'high risk' trials and deemed the risk of bias to be minimal as exclusion had very little effect on meta-analysis outcomes proving greater confidence in the results. The overall shift in results from the initial review to the update is likely to reflect recent developments in ante and post natal care and the growing awareness of the importance of post natal depression.

For maximum effect these types of interventions rely on national enrolment. Relying solely on a set of risk factors may not be sufficient to prevent a number of women from suffering from depression who do not necessarily meet these criteria. The review does not address the cost-effectiveness of psychosocial and psychological preventative measures at this scale and thus though seemingly effective, it may not be possible to implement these on a broader national scale.

Despite the existence of potentially preventative measures these are generally not employed in the UK National Health Service (NHS) and a more informal less aggressive approach to prevention is taken. NICE guidelines issued in 2007 recognise that preventative interventions are available and advocate some small intervention for pregnant women who have some symptoms of depression, but who do not meet the threshold for formal diagnosis, to prevent the development of serious depression. For women who fall into this category and also who have a previous history of depression or serious anxiety, NICE recommends brief psychological treatment comprising 4 to 6 sessions. This may take the form of either cognitive behavioural therapy (CBT) of interpersonal psychotherapy (IPT). For women who have not had a previous episode of depression or anxiety, social support during pregnancy and the post natal period through informal individual or group-based support is suggested. Psychosocial interventions such as those listed in the review outlined above designed to reduce the likelihood of developing mental health problems during pregnancy or the post natal period are deemed inappropriate for routine care. As a consequence of the lack of prevention effective treatment strategies need to be in place.

Far greater promise has been associated with treatment options, which have been deemed both clinically and cost effective (Hendrick, 2003, NICE, 2007, McCrone et al., 2004), but these can only be applied when cases are identified. Screening all new mothers with a well-established screening tool may improve case detection.

One of the main barriers faced by health professionals in the diagnosis of postnatal depression is the practicality of actually being able to identify the presence of symptoms in the limited consultation time available for a regular GP appointment which is generally around 10 minutes (Wilson et al., 2002). This issue is further compounded by the fact that postnatal visits with the GP are generally more focused on the baby's

development rather than on the mother's wellbeing. The more predominant barrier to case finding however is the hesitation of the mother to report her symptoms, often in fear of the consequences (Shakespeare et al., 2003, Poole et al., 2006). In the report by 4Children (2011) an overwhelming majority of mothers cited a lack of knowledge about the condition and fear of the stigma attached as the primary reasons they did not seek professional help with their symptoms. This presents obvious barriers to prompt and effective diagnosis and requires action at the societal level to increase recognition and awareness of the condition and provide reassurance to women that help will be provided. Given that in many cases depression will not spontaneously resolve (Sexton et al., 2012), the greater the delay from the onset of postnatal depression to the beginning of adequate intervention, the longer the duration the depressive period is likely to be maintained.

1.9 Treatment

Numerous studies have found anti-depressants effective in general depression particularly when combined with psychological based intervention such as cognitive behavioural therapy (CBT) (Stowe et al., 1995, Appleby et al., 1997, March et al., 2006). However, one major concern of mothers suffering from post natal depression is the effect of anti-depressants on children via breast feeding though scientifically this is becoming less of an issue due to progressions in drug manufacturing and the emergence of evidence that some anti-depressants can be used whilst breast feeding without posing significant risk to the baby (Gjerdingen, 2003). Despite these findings some women still have reservations about medication during breastfeeding thus effective psychological and psychosocial therapies are highly desirable (Pearlstein et al., 2006. Cognitive behavioural therapy (CBT) is one of the most commonly cited effective treatments for depression (Clark et al., 2003, Hendrick, 2003, Milgrom et al., 2005). CBT is a psychological approach that aims to resolve problems concerning dysfunctional emotions, behaviours and cognitions through a goal-oriented, systematic procedure. CBT aims to identify beliefs, feelings and behaviours associated with psychological disturbance and adapt them to a more positive line of thinking (Jackson et al., 2009). Individuals in receipt of CBT generally receive between five and 20 session on a weekly basis. Despite the large evidence base supporting the use of CBT, in the UK there is a large shortage of trained therapists in the field, meaning there can sometimes be delays

in actually being able to see a therapist (Proudfoot et al., 2003). More recently evidence has pointed towards the use of psychosocial or self-help methods of treatment. One example of a developing self-help measure is computerised cognitive behavioural therapy (CCBT). This is based on the same principles as standard CBT but can be self-administered by the patient in their own at their own convenient time with minimal therapist assistance. To date there is not a great wealth of evidence assessing whether or not CCBT is as effective as CBT specifically for postnatal depression but the limited data available has suggested CCBT may be both effective in treating symptoms (Kaltenthaler et al., 2002, 2008) and also cost effective (McCrone et al., 2004, Gerhards et al., 2010) but is susceptible to high fall out rates and low levels of adherence (Kaltenthaler et al., 2008). Given the circumstances of post natal depression and the time consuming process of raising a new born infant the low rates of participation are understandable. More research is required to fully assess the benefit of self-help methods for post natal depression and methods of making them more accessible and time efficient.

Psychosocial treatment options currently hold more promise than self-help measures. Telephone based peer support is one good example of this (Dennis, 2003). A randomised controlled trial was carried out which randomly assigned women to either usual care or peer support. Women who had previously experienced post natal depression themselves were trained in delivering telephone based counselling to mothers scoring above a defined threshold for depression on the EPDS (>12 points). Frequency of calls was not standardised and was suited to the requirements of each participant. A minimum of 5 calls was made to each mother with a mean duration of 34.4 minutes per call. Significantly more mothers in the experimental group exhibited decreased symptoms of depression at the first follow-up session (4-weeks after treatment commenced) (X^2= 5.18, df = 1, p = .02) and the final follow-up session (8-weeks post treatment onset) (X^2 = 6.37, df = 1, p =.01) assessment. At the 4-week follow up only 10% of women in the treatment group scored above threshold on the EPDS compared to 40.9% of mothers in the control group. Comparable findings were found at the 8-week assessment, at which time 15% of the treatment group scored above threshold compared to 52.4% of mothers in the control group. The results indicated that the peer-support intervention significantly decreased depressive symptoms at the 8-week assessment where mothers who received support were almost 5 times more likely

to have recovered from depression symptoms compared to controls (OR = 4.7; 95% CI, 0.91 to 25.46).

Home listening visits are one further alternative to structured psychological therapy. These are generally provided to individuals by community health visitors and involve reflective listening, a counselling technique in which the listener responds to verbal and non-verbal messages and reflects these back to the individual, and collaborative problem solving in which the pair work together to draw up lists of problems and potential solutions (Segre et al., 2010).

As a consequence of the numerous treatment options for postnatal depression, a systematic review published by the Cochrane Collaboration, was undertaken to drawn comparisons between the different options (Dennis and Hodnett, 2009). Studies were included that had identified depression within the first year after birth using self-report measures (EPDS or BDI) or through diagnostic interviews. The types of interventions included varied from psychosocial interventions such as counselling, listening based home visits and peer support, to more formal psychological interventions such as CBT, IPT and PDT. These interventions were compared to usual care which refers to any appropriate medical care received included drug therapy (antidepressants). Ten trials were included, nine of which provided useable data for 956 women. Five trials were conducted in the UK, with the remainder divided between the US, Canada, Australia and Sweden. Comparisons were made between psychosocial and psychological interventions both with usual care and with each other. Trials selecting participants based on a clinical diagnosis of depression were just as effective in decreasing depressive symptomatology as those that enrolled women who met inclusion criteria based on self-reported depressive symptomatology. Meta-analytic results revealed that any type of psychosocial or psychological treatment, compared to usual care, was associated with a reduction in the likelihood of continued depression at the completion of treatment, relative risk (RR) = 0.68, 95% confidence intervals (CI) 0.58-0.79. Psychosocial treatments alone were effective in decreasing depressive symptoms (four trials, n =231; RR = 0.49, 95% CI 0.28 to 0.85). The most beneficial of these treatments was non-directive counselling (three trials, n = 189; RR = 0.55, 95% CI 0.32 to 0.94). Beneficial results were found when assessments were completed immediately post-treatment for cognitive behavioural therapy (four trials, n = 209; RR = 0.79, 95% CI

0.62 to 1.01), interpersonal psychotherapy (one trial, n =120; RR =0.80, 95%CI 0.66 to 0.98), and psychodynamic therapy (one trial, n = 95; RR = 0.48, 95% CI 0.29 to 0.80). All treatment measures were shown to still be beneficial at the end of the follow up period (varied between studies, max 12 months) with the exception of PDT which was no longer beneficial at 9 months after birth. With the exception of PDT the meta-analytical results suggest that both psychosocial and psychological interventions are effective treatment options for women suffering from postpartum depression. However, none of these studies followed women up beyond the first year after birth thus the long-term durability of treatment remains unclear.

Despite their being a wealth of potential treatment options available for postnatal depression, treatments provided within the UK health service will likely be those recommended by NICE guidelines. Given the potential for transmission of drug based therapies (antidepressants) through breast milk in breastfeeding mothers these are not recommended. Instead, home listening visits, self-help strategies such as CCBT or structured psychological therapy such as CBT as described above, are recommended as treatment options for women with postnatal depression.

In order to receive adequate treatment, women must first be identified. As discussed earlier in section 1.6, depression is an underdiagnosed condition, with health professionals often failing to identify women and women failing to identify themselves (Hearn et al., 1998). Screening for postnatal depression in mothers and providing adequate treatment is one potential means of intervention which could lead to fewer symptoms in women and improved outcomes for children.

1.10 Screening

In a classic screening programme a test is applied to apparently healthy volunteers in order to identify those individuals at high risk of having an otherwise unrecognised illness (Grimes et al., 2002). Probable cases can then be referred to their own doctor to confirm the diagnosis and for any required treatments. According to the National Screening Committee a series of 23 criteria relating to the condition the screen identifies, the test itself, the appropriate treatment and the proposed screening

programme must be met for a screening strategy to be implemented (UK National Screening Committee, 2009).

Current NICE guidelines recommend the use of two brief case finding questions (Whooley et al., 1997) to identify possible postnatal depression (PND); "Over the past month, have you been bothered by feeling down, depressed or hopeless?" and "Over the past month, have you been bothered by having little interest or pleasure in doing things?". If either answer is positive, this prompts further screening with more in-depth self-report measures such as the Edinburgh postnatal depression scale or the patient health questionnaire (PHQ-9) as part of subsequent assessment (NICE, 2007). Despite not being specifically designed to detect the presence of postnatal depression, comparative studies show high concordance between the EPDS and the PHQ-9 at 83% (Yawn et al., 2009). The PHQ-9 has also shown good sensitivity and specificity at detecting cases of depression in a postnatal sample (Gjerdingen et al., 2009). Either measure is thus likely to be appropriate for case-finding. The Whooley questions were initially employed without the backing of a solid research base, with the only empirical support for their use coming from two unrelated studies (Whooley et al., 1997, Arroll et al., 2005). Whilst these studies verified the efficacy of the Whooley questions in their ability to detect some mental health problems, these studies did not focus on a postnatal population. Since their implementation, the Whooley questions have been validated in a specific maternal mental health context and have been found to have high sensitivity (100%, 95% CI 77%−100%) and moderate specificity (68%, 95% CI 58%−76%) (Mann et al., 2012). However it is still uncertain whether the Whooley questions are the optimal case finding strategy as women may not be aware of their depression or recognise that their symptoms are not normal. Questionnaires asking about specific symptoms such as persistent feelings of self-blame that are not obvious signs of depression may still be more useful means of case detection. Additionally, despite being recommended by NICE, a survey carried out by the Royal College of Midwives shows that a quarter of postnatal women are still not being asked about their mental health during routine appointments and around a third of midwives report that they do not feel they have the ability to recognise signs of depression in new mothers (RCM, 2014) suggesting that although some level of screening through the Whooley questions is advocated it is not universally adopted.

1.11 Cost-effectiveness of screening

When considering the implementation of a screening programme it is not only the efficacy of the screen that is important but also the associated costs. Screening strategies must thus provide a sufficient improvement to health within a reasonable budget. Whilst formal screening strategies for postnatal depression have been advocated, they are not currently considered cost-effective (Shakespeare, 2005, Hewitt et al., 2009).

The most recent evaluation of the cost-effectiveness of screening for postnatal depression is that conducted by Hewitt et al., (2009) in a HTA addressing methods for identifying postnatal depression in primary care. This report employed decision modelling to estimate the costs and health outcomes for a range of known identification strategies, such as the Edinburgh postnatal depression scale (EPDS, Cox et al., 1987) and the Whooley questions, over a 1 year time frame. Based solely on maternal outcomes, no screening measure, including the currently recommended Whooley questions, met a reasonable cost-effectiveness threshold, though the EPDS at its highest cut-off point was the closest of all the measures examined. This study was hindered by a lack of longitudinal evidence concerning the long-term outcomes associated with postnatal depression and more importantly the effects exposure to maternal depression has on children. It is possible that when accounting for the long-term negative impacts maternal postnatal depression has on children, screening for postnatal depression may be considered cost-effective.

This thesis thus aims to explore the long-term effects of maternal postnatal depression on child outcomes and develop further the decision model outlined in the HTA by Hewitt et al. (2009) to readdress the question of whether screening could be considered cost-effective when child outcomes are included.

1.12 Aims and Objectives

1.12.1 Aim

The primary aim of this thesis is to explore the effects of undiagnosed postnatal depression on child development over time and to address whether screening for

maternal postnatal depression could be considered cost-effective if child outcomes are included.

1.12.2 Objectives

Each chapter will have individual objectives. As a whole, the work carried out in this thesis has the following objectives:

- To explore the effects of undiagnosed maternal depression on child outcomes over time.
- To explore the individual pathways of maternal depression over time and how these affect child outcomes.
- To evaluate whether screening for postnatal depression might be cost effective taking into account the effects of undiagnosed and untreated depression on children.

1.13 Outline of thesis

Chapter 2   describes theory that attempts to explain why the early years are important and how effects may come about. It also describes the likely reasoning behind why maternal mental health appears to be so important for child development.

Chapter 3   discusses the key variables used throughout this thesis and presents the results from a review of reviews to identify the important covariates for building a statistical model to explore the effects on children of missed maternal depression.

Chapter 4   presents a statistical model exploring the effect of missed maternal depression on child outcomes up to age 7 using the millennium cohort study.

Chapter 5   explores the effect of maternal depression over time on child developmental outcomes.

Chapter 6   presents the results of a scoping review linking early childhood behavioural problems with long-term outcomes including education, criminal activity and labour market activity and earnings.

Chapter 7   builds upon an existing decision model evaluating whether screening for postnatal depression may be cost-effective after accounting for the effects on children of failing to diagnose and treat depression.

Chapter 8   provides an integrated discussion of the thesis findings grounded in the context of relevant theory and prior research. Implications for policy and clinical practice are outlined and suggestions for future research in the area are made.

Chapter 2: Theoretical underpinnings

2.1 Introduction

As described in Chapter 1, maternal postnatal depression can have lasting effects on child development. This chapter aims to describe why this is the case using combined principles from attachment theory and the life course approach to development.

2.2 Attachment theory

Attachment refers to the 'strong emotional bond that forms between infant and caregiver during the child's first year' (Hetherington et al., 2006, p238). Historically there have been several theories that offer explanations of how and why attachments develop. Freud's early psychoanalytic theory suggests that infants first become attached to the mother's breast and then to the mother herself as a source of oral gratification (Eagle, 1995). Learning theory postulates that infants become attached to their mothers because a mother provides food, or primary reinforcement (Bretherton, 1985). However, it is Bowlby's theory of attachment, which is grounded in developmental psychology and ethological theory that has become the dominant approach to understanding children's early social development (Misri & Kendrick, 2008). Ethological theory centres on survival behaviours (Bretherton, 1992) and was first applied to children in the 1960s. According to the ethological viewpoint, babies are biologically prepared to contribute actively to establish a bond with their caregivers, promoting the chances for survival. Aspects of children's social behaviour, including emotional expressions, cooperation, and social play, are thought to contribute to this (Pendry, 1998).

Attachment theory is built on the notion that instinctual infant and parent responses are important for survival and are thought to be biologically programmed to encourage a strong parent-infant mutualistic bond (Thompson, 2006). Infants are thus viewed to see their adult caregivers as safe and protective figures who demonstrate an active investment in the infant's survival. Attachments are further characterised by the tendency of infants to use their caregivers as secure bases from which they can

confidently explore their environments and engage in normal activities, a key aspect of cognitive development (Byng-Hall, 2005, Waters & Cummings, 2000).

The most important facet of attachment theory lies in its emphasis on the active role played by the infant's early social signalling systems through behaviours such as crying (Prior and Glaser, 2006) and the ability of the caregiver to respond to the infant's needs appropriately. From this perspective, attachment is viewed as a reciprocal relationship (Ainsworth, 1969).

The early notions of attachment theory still hold currency in contemporary developmental psychology and are consistent with current theory and findings from research in cognitive neuroscience (Bretherton & Munholland, 2008).

2.2.1 Attachment quality and its consequences

Building on Bowlby's early notion of attachment theory, Mary Ainsworth devised an empirical measurement of attachment through a caregiver-infant task known as 'The strange situation' (Ainsworth et al., 1978) which classifies infants as either 'securely attached' or as one of three subtypes of 'insecure attachment', namely insecure-anxious/ambivalent, insecure-anxious/avoidant, and insecure-disoriented/disorganized. In this task the caregiver and the child are placed in an unfamiliar playroom whilst a researcher records specific behaviours as the child experiences separation from/reunion with the carer and the presence of an unfamiliar stranger. It is using this measure that researchers have been able to classify children and observe the effects of different styles of attachment over time.

Several studies have noted that attachment can shape personality development (Sroufe, 2005, Weinfield et al., 2000, Cicchetti et al., 2006) and can also be a risk factor for later behavioural problems. These can persist to school age leading to poor peer interactions and higher parent and teacher ratings of problematic behaviour (Lyons-Ruth et al., 1996, 1999, Green et al., 2000), and to adolescence as research shows that adolescents who had poor attachment with their primary caregiver during infancy have higher levels of overall mental health problems at 17 years of age (Carlson, 1998).

In addition to behavioural development, interactions with attachment figures can also shape children's cognitive and social skills (Thompson, 2006, Stams et al., 2002). Securely attached children have been reported to be more socially competent, more peer-oriented and less dependent on adults than insecurely attached children by age 12 compared to controls (Carlson et al., 2004). Such effects remained present at age 19 when the socio-emotional functioning of these adults was also rated higher than controls.

2.2.2 Maternal depression and mother-infant attachment

Researchers have theorised for some time that reduced attachment quality is the driving force behind the relationship observed between impaired maternal mental health and child outcomes (Murray and Cooper, 1997). Secure attachment relies on effective parenting and from birth infants engage socially with their caregivers and are sensitive to the quality of their communication (Wan et al., 2009). The presence of postnatal depression causes functional impairment at a time when the actions carried out by the mother are vital to her infant's growth and development (Stewart, 2007) and is associated with reduced parenting responsiveness, affection and interaction (Goodman and Gotlib, 1999, Broth et al. 2004). Depressed mothers tend to report feeling less attached to and more negative toward their children compared with mothers who are not depressed (Cornish et al., 2006, Nagata et al., 2003, Rogosch et al., 2004). These alterations in parenting behaviour have been shown to have an adverse impact on the psychosocial and neurobiological development of their children (Newport et al., 2002).

Attachment disrupted by maternal depression likely creates a mechanism by which children are at risk of developmental disturbances, eventually leading to increased risk of developmental problems throughout childhood (Feldman, 2007, Murray & Cooper, 1997). Impaired maternal mental health can largely influence normal patterns of child development through disruption to the home environment and care giving to the infant, predominantly manifesting as a disruption in 'normal' parent-child interactions that would lead to the formation of a successful attachment (Murray et al., 1996, 1997, Field, 2010). Results from an in depth meta-analysis of 46 observational studies found a moderate association between maternal depression and negative and disengaged parenting behaviours (Lovejoy et al., 2000). The mother may find it difficult to maintain

focus on her child's development and requirements, and instead remain preoccupied with her own negative feelings causing her to miss her infant's cues and appear withdrawn and disengaged (Murray et al., 2003). Findings from another meta-analytical study also emphasise the role of maternal depression on attachment, collating data from American (n=5) and British (n=2) studies that compared groups of mothers with and without clinically diagnosed depression and assessed the attachment category of their infants using Ainsworth's strange situation (Martins and Gaffan, 2000). Results showed a significantly reduced likelihood of secure attachment compared to controls (p=0.0034) with studies showing a mean of 53% of children from depressed mothers having insecure attachment types, though some studies reported figures as high as 80% (Teti et al., 1995).

An important influence on a child's early cognitive development is the provision of a variety of activities and opportunities for play and conversation, as well as responsiveness to the child's signals and the teaching of specific skills (Ginsburg, 2007). Observational research has noted that mothers with depression communicate with their children less (Currie and Rademacher, 2004) and are much less responsive to their needs (Milgrom et al., 2004). Infants thus receive less feedback in response to their behaviours and can in turn become less responsive themselves (Diego et al., 2009). Depression is also sometimes associated with intrusive and even hostile play when the mother may fail to recognise the baby's discomfort and persist in trying to gain the baby's attention often by over stimulating (Murray et al., 2003). This may directly cause infant distress and deregulated behaviour in addition to having long-term effects on behavioural development (Morrell and Murray, 2003). Conversely depressed mothers may engage in less play based activity with their children than non-depressed mothers which can further impede development (Field et al., 2010).

Some of the potential consequences of insecure mother-infant attachments include externalising behaviour problems (Madigan et al., 2007) as well as internalising problems (Madigan et al., 2013) and impaired language development (Lemche et al., 2007) during early childhood. Attachment problems can also have long-term repercussions and appear to be related to conduct problems in preschool and later school periods, conflict with peers and parents, poor modulation of affect and increased risk of psychiatric problems (Gelfand et al., 1990, Granot et al., 2001).

Attachment theory offers one explanation of how maternal mental health can influence child outcomes, emphasising the importance of the interaction between the mother and the child as well as the timing of secure attachment formation. The life course approach builds on this theory, likewise emphasising the importance of timing, but also describing the effects of accumulation of negative events, in this case exposure to impaired maternal mental health.

2.3 The life course approach

The Life Course Approach to health arose in the 1980's as a consequence of novel data emerging from longitudinal studies that had not been previously available (Bartley, 2004), and is the current dominating perspective describing how early child circumstances affect long-term development. More recent longitudinal studies such as UK cohorts with prospective social and biological information collected in childhood and adulthood provide particularly valuable sources of life course data and have further shaped the life course approach. For instance the 1958 UK birth cohort study provides several examples of possible psychosocial processes that link childhood adversity to adult health and economic outcomes (Power et al., 2002, Montgomery et al., 1996).

The key proposal is that outcomes in adulthood may be a consequence of complex combinations of social, psychological and biological circumstances taking place over time from early life (Davey-Smith et al., 2002), with the main focus being concerned with the effects on health and health related outcomes of biology (including genetics), the environment and also social exposures during gestation, infancy, childhood, adolescence and adulthood (Kuh et al., 2003). As the life course is continuous, a wide range of exposures are encountered over time. Such exposures in later life may interact with exposures in early life either to enhance effects (synergism) or diminish them (antagonism) with timing and duration of exposure often mitigating effects (Kuh and Ben-Shlomo, 2004).

Life course outcomes describe a set of health, social adjustment, behaviour (including crime and substance abuse), educational attainment, and employment outcomes that occur during an individual's pathway into adulthood and emphasise the importance of

exposures throughout the life course, though most notably during gestation and early childhood (Kuh and Ben-Shlomo, 1997). Although individuals are more susceptible to exposures at certain periods in time, development continues to take place throughout life and consideration must be given to the numerous potential causal factors that manifest over the life course and variable importance of such factors at different points in time (White et al., 2005).

At all stages of the life course behaviours of family members seem to influence health and health related behaviour, though exposure to family members changes over the life course, meaning they have differential impacts on development at different time points. This is most notable during foetal development when maternal health and behaviour influence foetal growth and development in ways that can have lasting effects (Weinstock, 2005). In contrast, once children start school their contact with their parents is lessened and other figures such as teachers and peers begin to influence development.

There are three potential models that comprise the life course approach: the critical periods model, the pathways model and the accumulation model. These separate out the ideas of timing and quantity and introduce the notion that certain exposures can initiate developmental pathways. Within the life course approach there is considerable overlap between the three models and they not always distinguished by researchers in the same way (Kuh ans Ben-Shlomo, 2003). This is particularly true for the pathways and accumulation models which both focus on events over time. Models are not mutually exclusive and should not be viewed as competing theoretical perspectives but may operate simultaneously. As a result it is difficult to distinguish models empirically and to develop standardised and acceptable methods of combining cumulative exposure.

2.3.1 The critical periods model

The critical periods approach is drawn from biology and the idea that certain things will happen at particular times of development. A critical period of development thus refers to an often short space of time when development is rapidly occurring and individuals are particularly sensitive to the effects of exposures, i.e. a child can be particularly sensitive to a stimulus that would not have the same effect if encountered beyond this critical period. During these sensitive periods exposures to certain circumstances can

have lasting effects on development irrespective of later experience (Kuh et al., 2003), potentially due to impairments in brain development.

One of the most well-known examples that stems from the critical periods model is the concept of biological programming or latency modelling which is the basis of the foetal origins of adult disease hypothesis (Barker, 1998). Barker hypothesised that when biological processes occurring before birth are disrupted, for instance due to poor nutrient supply to the foetus, this can lead to the disruption and curtailment of growth. A theoretical example of this is the idea that nutritional disadvantage in utero followed by rapid growth in childhood is an underlying cause of adult risk of heart disease and diabetes. This implies that there is a critical period pre-birth in which adequate nutrition is essential for proper growth. Growth during early life has also been empirically linked to many other life course outcomes. Individuals who suffered growth failure during early life were found to complete less schooling and score lower on tests of cognitive skill in adulthood (Hoddinott et al., 2013). Although we cannot establish a direct cause and effect between these factors the association still merits emphasis on the importance of this critical period of development.

Different aspects of development are subject to different critical periods. For instance, the critical period for the development of binocular vision is estimated to be between three and eight months, though children remain sensitive until at least 3 years old (Fawcett et al., 2005). Researchers have also suggested critical periods for developmental processes such as language acquisition. This is a relatively long critical period thought to extend from infancy into puberty where it becomes increasingly difficult to learn language without specialist support (Hetherington et al., 2006). Galobardes et al. (2004) suggests that critical periods for psychological illness such as insecurity and instability exist in infancy and childhood and are affected by parental neglect. Sylva (1997) identifies that children are sensitive to care taking at ages six-eight months when they must develop core attachment to parents and ages twelve-thirteen months as being a critical period of intellectual and linguistic development. Critical periods can even occur before birth. The pre-natal period is considered a critical period for exposure to other potentially harmful circumstances such as maternal smoking, alcohol consumption and experience of infection (Eskenazi and Bergmann,

1995, Spohr et al., 1993). These can not only affect growth but also cognitive development and educational attainment (Horwood et al., 1998).

In relation to the present work the most important theoretical critical period is that for attachment. Bowlby proposed that the critical period of attachment formation is during the first 2 years of life where disruptions to the bonding process can lead to lasting problems, however much greater focus is applied to the importance of the first year (Bowlby 1969, Ainsworth, 1979). The importance of this early stage of life has not been challenged by modern research.

As maternal depression can impact on mother-infant interaction, depression during this critical period could have long-term effects. As shown in Chapter 1, postnatal depression occurring during the first postnatal year can have lasting effects on children's development (Grace et al., 2003, Goodman et al., 2011). Parenting quality, parental distress and the parental-offspring relationship in the early years can influence development of a range of psycho-social factors (Vondra et al., 2001, Sroufe, 2005). Hamilton and Hango (2008) examined the association between maternal psychological distress during children's early development and subsequent levels of distress in their offspring when they are in adulthood (aged thirty) as well as the influence of maternal characteristics at birth, using data from the BCS1970. Results indicate that greater symptoms of maternal psychological distress during early childhood are associated with greater symptoms of distress when the child reaches adulthood.

Biological based evidence grounded in neuroscience has also reported lasting effects of exposure to maternal stress during infancy. Essex et al. (2002) show that exposure to maternal stress during infancy can affect children's brain development resulting in heightened cortisol responses (a hormone produced in relation to stress) to stressors several years later in the preschool period. Such studies examining the effects of postnatal depression on child outcomes could be considered supportive of the critical periods approach as exposure during this time point can have lasting effects on development.

One of the objectives of this thesis is to examine whether the first postnatal year serves as a critical period for intervention targeted at maternal depression as a means of

preventing the potential detrimental effects on child development. In order to test this, comparisons will be made between the outcomes of children of depressed mothers who were and were not diagnosed and treated for depression in the first year after birth.

2.3.2 The pathways model

In the 'pathways model', poor childhood circumstances are seen as being the first stage on the pathway to poor adult health through providing influence on social trajectories. Advantageous or disadvantageous events may affect the risk of an outcome, such as disease, because they increase the probability of some other causal factor. This model is also known as a 'chain of risk' model considering exposures as a sequence of linked events where one negative experience tends to lead to another (Kuh et al., 2003).

Early childhood environments, particularly the family context which mediates the child's educational opportunities (Wadsworth, 1991) and the resources available within them, can place individuals on life trajectories that can influence well-being over time and set individuals on a pathway of lifestyle factors significant to life course outcomes. A consequence of reduced resources, such as family income and parent emotional status, in early childhood is often the development of problem behaviours in later childhood and adolescence (Conger et al., 1992). This pathways model emphasises the adolescent to adult transition through education as being particularly important because of its association with health related habits such as smoking exercise and dietary choice (Wadsworth, 1996).

Although the pathways model holds some intuitive value it is difficult to assess on an empirical basis. Longitudinal research is best placed to evaluate this theory, however such studies often only collect data at limited and often irregular time points which does not allow for continuous pathway effects to be confidently established. Nevertheless, there is some evidence to support the pathways approach. Feinstein and Bynner (2004) found that being in the lowest quartile of cognitive assessment scores at age five can be related to higher odds of having low income, having low qualifications and experiencing unemployment at age thirty (data from the 1970 British Cohort Study). Similarly Batty et al. (2007) also used data from BCS1970 from participants who had complete data on IQ scores at ages five and ten and risk factors at age thirty. They

sought to investigate whether a potential pathway exists in which early life mental ability influences adult risk factors for premature mortality where the mortality gradient included smoking, obesity, weight related diabetes, and raised blood pressure. Medical data were collated with information drawn from interviews at the age thirty follow up. Combining self-reported data with medical data in this way is a good way to reduce subjectivity associated with primarily self-reported data without losing the qualitative element. Results indicated that IQ at age ten was associated with socio-economic position, employment type, qualifications at thirty, and annual earnings by the age of thirty. Most risk factors for premature mortality were strongly associated with socio-economic circumstances such as smoking, being overweight or obese and self-reported high blood pressure, probably due to better education concerning these factors. This suggests that having a high childhood cognitive ability is beneficial to future health outcomes in adult life. Although it is not certain that these early circumstances are the sole cause of these outcomes it may be the case that they made individuals more susceptible to certain life events that led to the outcome.

The pathways model is similar to the critical period model in that it assumes that events can have lasting effects over time and impact upon future outcomes. It also has similarities with the accumulation model in that it suggests that events can increase the probability of another causal factor thus forming a cumulative effect.

2.3.3 The accumulation model

The accumulation model emphasises the importance of risks to health and well-being outcomes throughout the life course and suggests that exposures gradually accumulate over time although this does not preclude factors acting during critical periods having a greater impact (Kuh and Ben-Shlomo, 2004).

Focus is placed upon how both hazards and advantages accumulate over the life course and create inequalities by determining the risk of health and wellbeing outcomes. It identifies two main processes by which health and social status reinforce each other across the life course: social mobility and social protection. Social mobility means that parental socio-economic status influences many circumstances of childhood which in turn influence chance and direction of mobility into a different social class. This then

influences the accumulation of further health and social advantage and disadvantage. Social protection proposes that previous socio-economic circumstances can condition the impact of new disadvantage minimising the impact among the advantages and amplifying it among disadvantage (Blane, 1999). The accumulation of risk model suggests that effects accumulate over the life course although it also allows for developmental periods during which susceptibility may be greater (Galobardes et al. 2004). The accumulation of risk model can in itself follow a pathway. For example, children from families with lower income are more likely to have lower birth weights, poor nutritional intake, be more exposed to passive smoking, and have fewer educational opportunities (Ben-Shlomo and Kuh, 2002). Exposures thus form the 'chain of risk' denoted in the pathways model, by occurring collectively and accumulating. As the number, duration and severity of exposures increases, there is thought to be *cumulative* damage. Risk exposures may be independent from each other (Figure 2.1) or clustered (Figure 2.2).

One of the common reasons for clustering is where exposures all relate to an individual's or family's socioeconomic position in society. Being in a family characterised by low socioeconomic position during childhood is associated with lower birth weight, fewer educational opportunities, particularly the provision of in-home learning materials, increased family stress and the increased likelihood of been exposed to smoking (Kuh and Ben-Shlomo, 2004). All of these are a result of being raised in a low socioeconomic family, but all add together to form a cumulative effect.

The pathways model, outlined in section 2.3.2 (The pathways model), is also sometimes described as a variation of the accumulation model and is referred to as a chain of risk model whereby sequences of linked exposures lead to impaired function and increased disease risk (Hertzman et al., 2001). Each exposure in a chain of risk may not only increase the risk of a subsequent exposure (in a probabilistic rather than deterministic way) but may also have an independent additive effect on later function or disease (Figure 2.3).

The final variation of an accumulation model is thought of as a 'trigger effect' and describes a chain of risk type model but where it is only the final link in the chain that has any marked effect on outcomes (Kuh and Ben-Shlomo, 2004) (Figure 2.4).

49
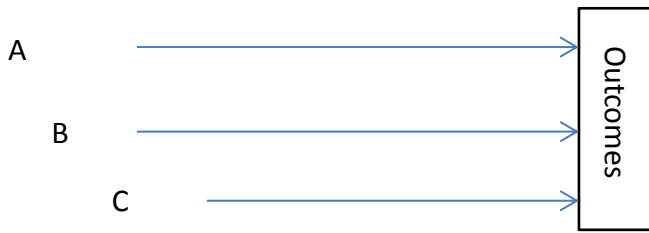
Figure 2.1. Independent risk (Kuh and Ben-Shlomo, 2004)
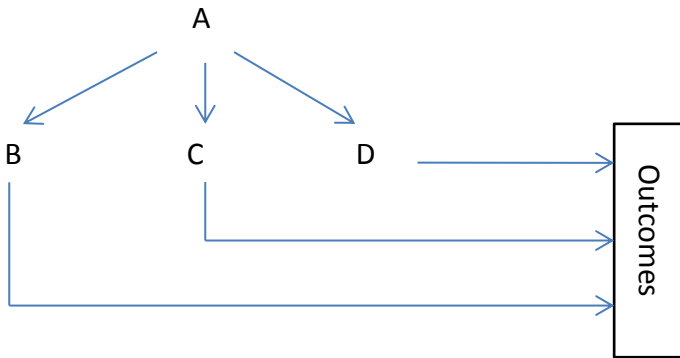
Figure 2.2. Clustered risk exposures (Kuh and Ben-Shlomo, 2004)
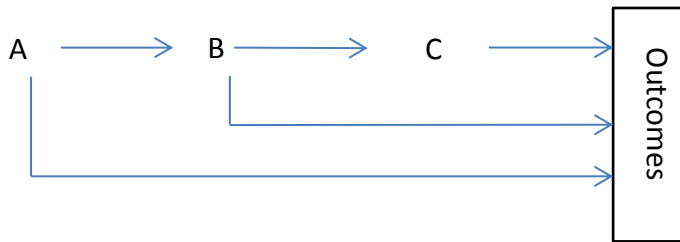
Figure 2.3. Chain of risk exposures (Kuh and Ben-Shlomo, 2004)
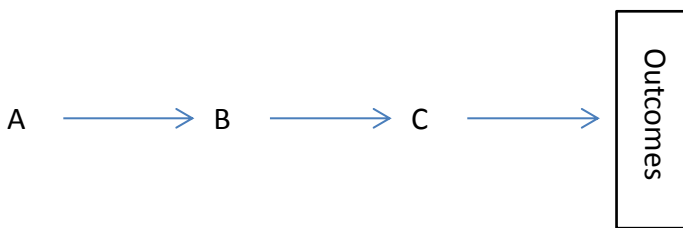
Figure 2.4. Alternative chain of risk model (trigger effect) (Kuh and Ben-Shlomo, 2004)

Although the assumptions of the accumulation model are intuitive, in that the more negative exposures the individual is exposed to the greater the effect on outcomes, testing these assumptions is difficult due to the need for in depth longitudinal data assessing multiple exposures over time (Ben-Shlomo and Kuh, 2002). However, there is empirical support from some longitudinal studies that have examined the cumulative effects of repeated exposures over time. For instance Najman et al. (2010) examine the

impact on mental health of repeated/prolonged exposure to poverty from early childhood to young adulthood finding that the more frequently the child was exposed to poverty, the greater the risk of anxiety and depression at ages 14 and 21 years.

Similarly, studies have explored the effects of persistent/ recurring depression on children's long term outcomes, again finding that the longer the duration of maternal depression or the more frequent the exposure to depressive episodes, the more affected the child's outcomes (Kiernan and Mensah, 2009, Ashman et al., 2008).

The accumulation model builds upon the critical periods model to take account of future exposures that could affect development. In relation to the questions addressed in the present work, this refers to further exposures to maternal depression across childhood.

2.4 Summary

This chapter describes the likely mechanism through which maternal depression affects children, namely the theory of attachment, as well as theoretical models that emphasise the importance of the early years and attempt to explain how childhood circumstances can influence life course outcomes and adult health, drawing upon evidence in the research literature.

Regardless of the model accepted, it is clear that circumstances in the early years of life are crucial for long-term development with negative exposures during this period having lasting negative consequences.

In the consideration of the relationship between maternal mental health and child outcomes for this thesis, there is a clear case for some overlap between the different life course models. Using the principles of the critical periods model alone, it could be hypothesised that poor mental health during the critical period for attachment, which is considered to be the first one to two years of life (Bowlby, 1969, Howard, 2011), could cause lasting detrimental effects irrespective of future circumstances. Conversely, the accumulation model takes into consideration experiences over the entire life course thus suggesting that future experiences, such as repeated/prolonged exposure to impaired

maternal mental health can hold a cumulative effect on development, changing the modifying the effects of any experiences confronted during the critical period.

Work carried out in this thesis will explore both the critical periods model and the accumulation model, examining how early exposure to maternal depression combined with future exposure affects cognitive and behavioural outcomes up to age seven.

Failing to diagnose and treat maternal depression during the first postnatal year could lead to more negative effects for children. If this is the case there may be support for the notion of screening for postnatal depression and may possibly reverse the decision presented in the HTA report conducted by Hewitt et al. (2009) discussed in Chapter 1.

2.5 Hypotheses

From these models 2 key hypotheses were generated which will be explored through the research conducted in this thesis:

1. There is a critical period of development during the first postnatal year where exposure to maternal depression can have lasting effects. Treating maternal depression may protect against these effects, thus there may be a critical period for treating maternal depression in relation to child outcomes.

2. There may be a cumulative effect in addition to the critical period where effects on development may be accentuated by future exposures.

2.5.1 Testing the hypotheses

The critical periods model will be tested using multilevel modelling comparing children of diagnosed and undiagnosed mothers with each other and with non-depressed women. Depression status will be assessed during the first postnatal year only and child outcomes will be assessed up to age seven in line with availability of MCS data.

Accumulation of depression will be tested using multivariate regression models assessing the effects on child outcomes at age 7 (again in line with MCS data) according to trajectories of mental health over time.

Chapter 3: Model development

3.1 Introduction

Chapters 1 and 2 describe the importance of maternal mental health and the theoretical models that attempt to describe how early years circumstances can affect long-term outcomes.

The primary aim of this thesis is to explore the effects of undiagnosed postnatal depression on child development over time and to address whether screening for maternal postnatal depression could be considered cost-effective if child outcomes are included.

Differences in child outcomes will be examined firstly according to differences in maternal mental health and diagnosis during the postnatal year in line with the critical periods model. As described in Chapter 2, future exposures to negative circumstances, such as poor maternal mental health, can enhance those effects inflicted during the critical period, known as a cumulative effect. This forms the basis of the accumulation model which will also be tested by exploring the effects in changes in maternal mental health over time and how these impact upon child developmental outcomes. In this sense a chain of risk exposure model is thought to be taking place (Figure 2.3), whereby each exposure in a chain of risk can increase the risk of a subsequent exposure, in this instance an episode of maternal depression could be predictive of a future episode of maternal depression (Josefsson et al., 2007), with each episode potentially having either an individual or cumulative effect on later development.

Statistical modelling will be used to explore these ideas, comparing children of undiagnosed and untreated depressed mothers with children of diagnosed and treated depressed mothers and children of non-depressed mothers. This chapter aims to discuss the source of data and the variables that will be used in the model.

3.2 Data sources and outcome measures

3.2.1 The Millennium cohort study

Studies examining the influence of parental mental health on offspring have been hindered by a shortage of longitudinal datasets that contain prospective data on parents and their children over numerous decades. Such studies can tell us how early disturbance to the mother's state of mind influences children over time. The MCS is a large-scale longitudinal survey of the families of babies born in the four constituent countries of the United Kingdom in 2000/1 (Dex and Joshi, 2005). A noteworthy attribute of the MCS is that it is disproportionately stratified by UK country and by type of electoral ward within country. There is also more than one respondent per household, however in this instance only the mother's responses are considered. At the time this research was conducted, data collection sweeps have been carried out when the cohort child was around nine months (MCS1), three years old (MCS2), five years old (MCS3) and seven years old (MCS4). For the majority of families, the mother acted as the primary responder, though data was also collected from fathers and other family members such as partners and grandparents where possible. Data were collected through face-to-face computer assisted personal interviews (CAPI) and computer aided self-completion interviews (CASI).

The MCS adopted a carefully constructed sampling design to include 19,517 children in 19,244 families acquired through UK Child Benefit Records and was designed to offer an accurate representation of the total population. In order to do this, particular sub-groups of the population were intentionally over-sampled, including those living in disadvantaged circumstances, children from minority ethnic backgrounds (in England), and children growing up in the smaller countries of the UK. Wards in which over 30% of the population were in the categories 'Black' or 'Asian' were classed as 'ethnic minority' wards. 'Disadvantaged' wards were those that did not meet the requirements to be considered as ethnic minority but which were included in the poorest 25% of all wards based on the Child Poverty Index. 'Non-disadvantaged' wards were those not meeting the criteria for either of these two wards. The majority of those wards meeting the criteria for ethnic minority wards would also have met the criteria for disadvantaged wards (Plewis et al., 2004). The complete overall response rate was 72%, however,

compared to the non-disadvantaged wards, response rates were lower for families residing in disadvantaged areas or ethnic minority areas. A more comprehensive overview of the MCS sampling strategy can be obtained in Plewis et al. (2004) and in the technical report on sampling (Plewis, 2007).

The benefit of the Millennium Cohort Study is its prospective and longitudinal nature which allows us to take snapshots at various time points and place data on a time continuum. The diversity of the MCS sample means that the data can be standardised to the majority of the UK population. In terms of the question at hand, the MCS can allow us to ascertain when disturbances commence and cease and also when are most problematic. Although there is support for the notion that maternal depression affects children, these effects are generally reported when individuals have actually been clinically diagnosed and provide little insight into the effects of undiagnosed depression that may have been missed by health care professionals.

3.2.2 Maternal postnatal depression

As discussed in Chapter 1, the primary aim of this thesis is to identify the effect of undiagnosed maternal depression, that is depression that is evident on self-report scales but has not been diagnosed and treated by a clinician, during the postnatal period on child outcomes over time.

As discussed in Section 1.4, there are two main guidance documents designed to guide clinicians in the diagnosis of depression, the ICD-10 and the DSM-V. However, there are many self-report measures that can also be used to identify women that are experiencing symptoms of depression.

3.2.2.1 Methods of self-report

The two most well-known self-report questionnaires often used as indicators of postnatal depression are the Edinburgh Postnatal Depression Scale (EPDS) and the Beck Depression Inventory (BDI), though the latter is also a measure of general depression.

The EPDS comprises 10 questions used to identify the presence of general depressive symptoms (Cox et al., 1987). Although widely used, the EPDS has been criticised for failing to take into account depression symptoms that are exclusive to postnatal depression (Harlbreich and Karkun, 2006). However, the EPDS performs well at identifying both major and minor episodes of depression and has demonstrated good sensitivity (0.60-0.86) and specificity (0.45) (Hewitt et al., 2009).

In contrast the BDI comprises 21 questions, over double that of the EPDS. The BDI-II was introduced in 1996 to reflect the criteria for depression within the DSM-IV and has demonstrated high internal consistency (Beck, et al., 1996). As the BDI was not designed to specifically detect postnatal depression, some of the questions are likely to be inappropriate for a postnatal sample. For example, the BDI asks questions concerning appetite and changes in libido and sleep, all of which are likely to be affected by child-birth and are not necessarily indicative of depression (Conrad et al., 2012). Consequently the use of the BDI is likely to result in a high proportion of false positive cases.

In addition to these measures, the PHQ-9 may also be used to detect the presence of depression. This 9-item self-administered questionnaire has demonstrated good sensitivity (0.88) and specificity (0.88) for detecting major depression (Kroenke, et al., 2001) with comparison based studies showing high concordance between the EPDS and the PHQ-9 at 83% (Yawn et al., 2009).

3.2.2.2. MCS measures

Malaise Inventory

In the MCS, mothers at sweep 1 (9 months) are asked a series of questions to determine their mental health, such as '*how often have you felt low or sad during the last 2 weeks?*'. As part of this series a reduced version of the Malaise Inventory (Rutter, 1970) was also asked of mothers. The Malaise Inventory is a 9 item self-completed scale whereby high levels of psychological distress have previously been categorised by a score of four or more points (Bartley et al., 2004, Mensah and Kiernan, 2010).

Questions such as *'do you often get worried about things?'* and *'do you often feel depressed?'* are presented in a yes or no format (see Box 3.1).

Box 3.1. Malaise Inventory items

Malaise Inventory	Yes	No
Do you feel tired most of the time?	☐	☐
Do you often feel miserable or depressed?	☐	☐
Do you often get worried about things?	☐	☐
Do you often get into a violent rage?	☐	☐
Do you often suddenly become scared for no good reason?	☐	☐
Are you easily upset or irritated?	☐	☐
Are you constantly keyed up and jittery?	☐	☐
Does every little thing get on your nerves and wear you out?	☐	☐
Does your heart often race like mad?	☐	☐

Traditionally the Malaise Inventory consists of a 24 item scale which has been broadly used as a measure of distress and depression levels (Grant et al., 1990, McGee et al., 1986) and has widely been shown to have good reliability and validity (Grant et al., 1990). The full Malaise Inventory has been used in both large scale surveys of samples drawn from general populations (Dodge and Silva, 1980) and in smaller more constrained sub-groups, e.g. mothers of severely disabled children (Hirst and Bradshaw, 1983). Although the full version of the Malaise Inventory has been rigorously tested, the MCS employs a reduced 9 item version of the scale which has not been robustly validated or tested against other more established measures such as the EPDS. However, the reduced measure has been used in other comparable literature (Bartley et al., 2004, Mensah and Kiernan, 2010, Kiernan and Huerta, 2008) and although it does not guarantee a diagnosis of depression, it can provide a good inference of the level of symptoms a mother is experiencing. Additionally, the correlation between Malaise scores and clinical diagnosis was assessed in the current sample and a moderate correlation was found ($r=0.3650$, $p<.0001$) indicating that the reduced Malaise

Inventory used in this research is likely to be a good predictor of depression. However, it is important to note that it is not the intention to derive a clinical diagnosis of mental illness from this self-report data, rather the intended analysis aims to recognise mental health symptoms which can be indicative of underlying mental illness as experienced by mothers who have specifically not received an official diagnosis of depression from a clinician. Throughout this thesis, women scoring above threshold on the Malaise Inventory will be deemed 'depressed'.

Kessler 6

At the following surveys (ages 3, 5 and 7 years) mothers complete the Kessler 6 scale via a computerised self-report whereby a score of at least seven points is indicative of depression. Unlike the Malaise Inventory which requires respondents to answer on a yes or no basis in regards to symptoms, the Kessler 6 allows respondents to rate their symptoms more in terms of severity with answers ranging from '*all of the time'* to 'none *of the time'* (Box 3.2).

Box 3.2. Kessler 6 items

Kessler 6 (K6) During the past 4 weeks (28 days), how much of the time did you feel….					
	All of the time	Most of the time	Some of the time	A little of the time	None of the time
so sad nothing could cheer you up?	☐	☐	☐	☐	☐
nervous?	☐	☐	☐	☐	☐
restless or fidgety?	☐	☐	☐	☐	☐
hopeless?	☐	☐	☐	☐	☐
that everything was an effort?	☐	☐	☐	☐	☐
worthless?	☐	☐	☐	☐	☐

It has demonstrated good internal consistency and reliability, and has consistent psychometric properties across major socio-demographic sub-samples (Kessler et al. 2002, Kessler et al. 2003). Like the Malaise Inventory the Kessler is not intended to provide a firm diagnosis but can give an indication of the presence and scale of symptoms.

3.2.3 Outcome measures

In Chapter 1, an update of a systematic review was undertaken to address the currency of the association between maternal depression and child outcomes and to identify recent literature. The results of this review showed cognitive and behavioural development to be particularly affected by maternal depression, thus these aspects are the focus of the current research.

3.2.3.1 Cognitive development

Of the papers identified in the systematic review presented in Chapter 1, seven explored the association between maternal depression and child cognitive development.

Two of these studies used measures from the Bayley Scales of Infant Development (BSID) to address child cognitive development. This measurement is primarily used to assess the motor and language development (receptive and expressive) of infants and toddlers, ages 0-3 and consists of a series of developmental play tasks which are observed and scored (Black and Matula, 1999). Three further studies used measures from the Bracken Basic Concepts Scale (BBCS) which assesses children's comprehension of: colours, letters, numbers, sizes, comparisons of objects and shapes which provide an indication of the child's readiness for formal schooling (Bracken, 2002). These measures are similar in that they both address multiple facets of development reflecting the broad range of skills that fall under the heading of 'cognitive skills, and both measures are frequently used and have established validity (Antasi and Urbina, 1997, Bracken, 1984).

The two remaining studies used education attainment measures including GCSE scores (Murray et al., 2010), and results from the Foundation Stage Profile (Kiernan and Mensah, 2009) which provides information at both national and Local Authority level

on achievement outcomes at the end of the Early Years Foundation Stage, a statutory stage of the National Curriculum for England.

In addition to these assessments identified through the results of the systematic review, there are many other measures available for the assessment of child cognitive skills. One of the most prominent of these is the Wechsler tests, used predominantly in the United States but also in European based research including that conducted in the UK (Domino and Domino 2006). The Wechsler tests comprise of two intelligence tests, the Wechsler Intelligence Scale for Children (WISC) and the Wechsler Preschool and Primary Scale of Intelligence (WPPSI). The WISC-IV consists of 15 subtests, 10 of which are used to derive a full scale IQ. The subtests are divided into verbal (e.g. vocab, arithmetic, comprehension) and performance (e.g. picture arrangement, block design). The WISC–IV has been standardised in the US, based on a sample of 2,200 children between the ages of 6 and 16 years 11 months (Chen and Zhu, 2008), and in the UK, based on a sample of 780 children (Whitaker, 2010). The WPPSI-III is a modified version of the WISC-IV comprising 14 sub-tests and is targeted at 3-7 year olds. As in the WISC-IV, children are assessed in terms of their verbal and performance abilities from which a full IQ score can be derived. As with the WISC-IV, the WPPSI-III has been standardised in both a US and UK context with 1,700 children included in the US sample and 805 children included in the UK sample (Kaufman and Lichtenberger, 2000).

Though these measures are considered valid methods of assessing child cognitive development, in 1965 the British Psychological Society sought to develop a measure standardised on a British sample that would provide a profile of abilities rather than an overall IQ indicator. This led to the construction of the British Ability Scales (Elliot et al., 1979). It is this measure that was used in the Millennium Cohort Study and will form the basis for the analysis conducted in this thesis.

British Ability Scales; Language development

In the Millennium Cohort Study, subscales of the British Ability Scales (BAS) (Elliot, 1979), are used to assess language development as an indicator of cognitive functioning at the three, five and seven year sweeps.

The various subtests of the BAS have been shown to be significantly correlated with each of the others and can be combined to produce a General Conceptual Ability (GCA) score. This is carried out through the summation of scores from subscales evaluating verbal skills, spatial awareness skills and non-verbal reasoning (Connelly, 2013). These general scores, which are similar to IQ scores, have been shown to correlate highly with independent measures of school and academic attainment (Elliott, 1983). However, in the MCS participants are only asked to complete one or two of the available subscale at each sweep thus making the derivation of a general intelligence indicator not possible.

The BAS subscales included at each of the MCS follow up data sweeps varied, however, assessments of language were included at each sweep thus reflecting a usable repeated measure for the present analysis. At ages three and five language development was measured using the 'Naming Vocabulary' subscale, and at age seven was measured using the 'Word Reading' subscale of the British Ability Scales (Elliot, 1983).

The naming vocabulary subscale provides an estimate of children's expressive language skills. The child is asked to name a series of 26 pictures of everyday items such as *'jar'* and *'sink'*. The number of items shown to each child is dependent on their performance. If five successive items are answered incorrectly the questions are stopped. The BAS was administered in English to children who could speak English, and was also offered in Welsh. As opposed to the naming vocabulary assessment, the word reading assessment issued at age seven measures receptive language skills. In this instance children were presented with cards with words in text and were asked to read the word aloud. As expressive skills, which require a child to say something, were assessed in the naming vocabulary subscale, whilst receptive skills, which demonstrate an understanding of language, were assessed in the word reading subscale, the two scores are not strictly comparable. However, the correlation between the Word Reading score at age 7 and the Naming Vocabulary score at age 5 is 0.34 (Hansen et al., 2010) suggesting they are somewhat equivalent. Because the scores are not directly comparable the raw scores from each of the subscales were dichotomised in the present analysis to separate the lowest scoring 10% of children at each of the time points. This allowed for the assessment of whether exposure to maternal depression, diagnosed or undiagnosed, affected children's likelihood of being in the poorest performing 10% of children.

As information was only available for language development, the present analysis cannot fully address the impact of diagnosed and undiagnosed maternal postnatal depression on child cognitive development as a whole, but only on the specific facet of language development. Despite this, early vocabulary skills have been repeatedly highlighted as a crucial part of cognitive development (Gathercole et al., 1992, Ricketts et al., 2007, Joshi, 2005) that can be predictive of long-term educational attainment (Duncan et al., 2006), thus can provide valuable insight into how exposure to maternal depression may affect intelligence based outcomes.

The BAS is completed by the children themselves rather than the mother, thus reducing the likelihood of bias. This makes language development an objective way of measuring the effect of exposure to maternal depression on children.

3.2.3.2 Behavioural development

Twelve papers exploring the association between maternal depression and child behavioural development were identified in the systematic review presented in Chapter 1. Of these, eight (67%) used the Child Behavioural Checklist (CBCL) as the measure of behavioural development. The CBCL is one of the most commonly used measures, describing symptoms of emotional and behavioural disturbance over the past 6 months and can be rated by the parents of the child. The CBCL is empirically driven and comprises 138-items for use with 4-16 year olds and assesses a wide range of pathological behaviours (118 items) and the child's social competence (20 items) (Achenbach and Rescorla, 2001). With well-established normative data and standardised clinical cutoffs, the CBCL internalising, externalising, and total problems scales have demonstrated strong psychometric properties within clinical settings for discriminating between diagnosed populations and those with no problems (Achenbach et al., 1991, Chen et al., 1994, Seligman et al., 2004). Despite being highly correlated with clinical diagnosis (Ebesutani et al., 2010), the CBCL receives criticism for its length, for having an overly negative perspective and for being expensive to administer (Guttmannova et al., 2008).

One paper based in France (Melchior et al., 2012), used a French outcome measure, the EAS Questionnaire (Buss and Plomin, 1984), that has not been validated elsewhere and

shows only satisfactory psychometric properties in children and infants (Gasman et al., 2002).

The remaining three papers used the Strengths and Difficulties Questionnaire (SDQ) as a measure of child behavioural development. This measure was used in the Millennium Cohort Study and will form the basis for the analysis conducted in this thesis.

The strengths and difficulties questionnaire (SDQ)

Behavioural development was assessed in the Millennium Cohort Study at ages three, five and seven using the Strengths and Difficulties Questionnaire (SDQ) (Goodman, 1997) which measures four different constructs of behaviour: inattention/hyperactivity, conduct, peer relations, and emotional development. These individual constructs can be used to create externalising and internalising behaviour subscales. Two dimensions from the SDQ are used to make up the externalising behaviours construct: general conduct issues, and hyperactivity/ inattention, whilst the remaining two, namely peer related problems and emotional problems, are combined to form the internalising construct. Inevitably, this dichotomy into 'externalising' and 'internalising' is not perfect. For example, a child's internalising behaviour problems can have a negative impact on others, including siblings, parents, peers, and teachers, thus may appear as externalising symptoms. Similarly, children with externalising behaviour problems not only may negatively affect their outside world, but also may be psychologically suffering internally. In fact, there is significant and substantial co-morbidity between externalising and internalising behaviour problems (Hinshaw, 1987) meaning that it is important to study these constructs cumulatively. Despite these problems of definition and co-morbidity, it is both possible and useful to individually assess the constructs indicative of externalising and internalising behaviour problems as even in isolation the presence of these behaviours can have long lasting deleterious consequences. For example some authors have claimed that children with externalising behaviour problems are more likely to grow up to become delinquent as adolescents, and criminal and violent as adults (Simonoff et al., 2004, Murray et al., 2010, Klein et al., 2012). Similarly, children with internalising behaviour problems, such as emotional or peer-related problems, are more likely to develop depression and anxiety in later life (APA, 1994).

As well as assessing behaviours based on the externalising and internalising constructs total scores can be generated for children to take account of cumulative problems. The four scales which make up the externalising and internalising constructs are combined to yield the total behavioural difficulties score which is used to indicate whether the child is likely to have a significant behavioural problem.

Each of the questions in the SDQ are posed as statements such as '*considerate of other people's feelings*' or '*restless, overactive, cannot stay still for long*'. These are rated in turn by the mother using a scale from 0 to 2 whereby somewhat true is always scored as 1 but whether not true and certainly true are scored 0 or 2 depends on whether the item is framed as a strength or a difficulty. Responses were summed to provide a total score (max 10) for each dimension. A total score out of 40 can be generated by summing the scores of the four key sections; conduct, hyperactivity/inattention, emotional problems and peer problems, whereby a score of 0-13 is considered normal, 14-16 borderline and 17 and above abnormal. A score of seventeen or more is thus regarded as a high score, indicative of the child exhibiting behavioural difficulties outside the norm (Goodman, 1997 and 2001). However, the guidelines for use of the Strengths and Difficulties questionnaire (Youth in Mind, 2010) suggest that cut-off points should be adjusted in relation to the sample to ensure the highest scoring 10% of the population are identified as showing signs of difficulties. This scoring system has been standardised on a large sample of 10, 298 children aged 5-15 across the UK (Meltzet et al., 2000), which revealed a normal mean score of 8.4 points on the total difficulties scale and showed that 18% of children score within the borderline/ abnormal categories.

Due to the way in which this measure is administered through parental reports there is a possibility of bias. Some researchers have suggested that mothers with depression may be more likely than mothers without depression to report negatively on their children's behaviour and that the negative reports of depressed mothers are distorted by bias (Renouf and Kovacs, 1994). Other researchers have argued that there is no difference in accuracy between depressed and non-depressed mothers' perceptions but that negative maternal reports reflect a real difference in their children's problem behaviour (Boyle and Pickles, 1997 a, b). Richters (1992) carried out a review of the research literature concerning this association and found that there was no apparent evidence for distortion or bias in the reports of depressed mothers. In addition to this Richters also discovered

that some studies have actually suggested that depressed mothers can be more assertive and accurate in detecting behavioural problems in their children than non-depressed mothers (for example, Conrad and Hammen, 1989), and that mothers in general are often keener observers of their children's development and behavioural well-being than are other types of observers (Glascoe, 2005). In light of such findings it would appear that even when the SDQ is completed by depressed mothers it maintains its appropriateness as a tool for detecting likely cases of behavioural difficulties. It is important to note that the SDQ does not guarantee that a disorder is present in those individuals scoring high or that no disorder is present in those scoring low, but the instrument has proved useful for screening and case finding and is taken to be indicative of the key symptoms of behavioural difficulties. Furthermore, the SDQ is highly correlated with the CBCL, a more diagnostic measure, and has been shown to be equally as good at detecting internalising and externalising problems and significantly better at detecting inattention and hyperactivity problems (Goodman and Scott, 1999).

3.2.3.3 Covariates

As noted in Chapters 1 and 2, there are a range of different circumstances that may have an independent effect child development, such as poverty (Feinstein, 2003, Bhattacharya, 2010) and family structure (Manning and Lamb, 2003). Some of these factors can also increase the likelihood of a mother developing postnatal depression.

When developing models it is important to consider any potentially confounding variables, which are 'variables (that may or may not have been measured) other than the predictor variable that potentially affect the outcome variable' (Field, 2009, p783). To adequately identify the effects of one variable it is necessary to control for as many other potential confounding variables as possible. For this thesis potential confounding variables are those which may independently affect children's development in addition to maternal depression. As these variables are more likely to occur within the depressed sample, for example there are likely to be greater levels of poverty in the depressed mothers than in the non-depressed mothers (Liaw and Brooks-Gunn, 1994) it is important to include these variables in the model in order to ensure as far as possible that any differences observed are likely to be due to actual differences between groups and not error in the model.

Another important reason for accounting for these variables in any statistical models relates to accumulation theory. Accumulation does not have to occur over time through repeated exposures to the same negative stimuli. An accumulation effect can also occur when multiple negative stimuli are experienced at the same time (Ben-Shlomo and Kuh, 2002). It is thus important to account for as many of these potential confounding variables as possible to ensure a true effect.

3.3 Literature on potential confounders

Research focused on child development has identified many different circumstances that can affect child outcomes. In an attempt to summarise these circumstances, Shonkoff and Phillips (2000) reviewed the available literature describing the important contexts for child development, encompassing both family and child level characteristics and demographics in their book; 'From Neurons to Neighborhoods: The Science of Early Child development'. In this review a large emphasis is placed on the importance of nurturing relationships between the child and the caregiver (usually the mother), fostering secure attachments and providing a stimulating environment for development, with particular emphasis on maternal mental health, reinforcing the idea that maternal depression is highly problematic for child development.

They then went on to discuss other important contexts for child development, primarily focusing on socio-economic circumstances termed 'family resources', these included income and poverty, parental education, parental education, and family structure. These would need to be included in the current model to better isolate the effects of maternal depression on child development. Their findings are summarised below.

Poverty

Evidence was drawn from both experimental and non-experimental studies to highlight the importance of poverty for child outcomes, finding large effects of poverty on behaviour and associations between long-term economic status and child achievement (Bos et al., 1999, Gennetian and Miller, 2000, Morris and Michalopoulos, 2000, Brooks-Gunn and Duncan, 1997). It was also noted that young children are more likely than any other age group to live in poverty with family income generally rising over

time, however it is within these early years that negative circumstances have the greatest effects.

Parental education

From a theoretical perspective, parental education levels are known to be strongly related to the home literacy environment, parental teaching styles and investments in resources, all of which are able to promote learning. Parents with greater levels of qualifications are also more likely to supply high-quality child care, educational material, and take children on educational visits to libraries and museums.

Two studies were cited that looked at children of young mothers and differences between first and subsequent children when the mother has acquired further education between births. In the first one no effects of increasing maternal education were found on children's achievement scores (Kaestner and Corman, 1995), however the second one showed a moderate positive effect on children's vocabulary development per additional year of maternal schooling (Rosenzweig and Wolpin, 1994).

In these two studies the limited effects of education may be due to other co-occurring factors in addition to low levels of education associated with being a young mother such as increased risk of depression or living in poverty. It may also be the case that parental education at the birth of the first child shapes parenting styles and is not affected by subsequent increases in education.

Family structure

The review findings from Shonkoff and Phillips found that children raised by single-parents have lower levels of social and academic well-being than children from intact marriages. However, single parenting is commonly associated with both lower socio-economic status (McKay, 2003) and younger maternal age (Coley and Chase-Lansdale, 1998), both of which are independent risk factors for poorer child outcomes meaning that the effects may not be exclusively due to only having one care giver. As single parents generally survive on one income they may not provide as many stimulating resources as two parents could provide. Time constraints may also limit the amount of quality time single mothers are able to spend with their child; this may further hinder

their development through diminished emotional support and reduction in cognitive stimulation within the home environment.

Employment

Although parental employment is strongly linked to family poverty status it was also given independent consideration. Both maternal and paternal employment was noted to be important for shaping children's development. This was thought to be due to both income effects, mirroring those associated with poverty and also due to the decreased amount of time employed parents have available to devote to their children.

Evidence suggesting employment is important were said to be largely based on middle income families and thus not representative of other social groups. The findings presented were also inconclusive as to how important parental employment is for child outcomes. One study suggested that maternal employment in the first year has no negative effects and even a slight positive effect on children (Hoffman et al., 1999), thought to be due to mothers in employment having an increased sense of self-worth leading to better well-being and mental health.

In contrast other studies have shown that maternal employment during early childhood that involves long hours and thus prolonged separation from the child could be detrimental (Han et al., 2000, Waldfogel et al., 2000).

Shonkoff and Phillips (2000) concluded that the effects of employment are dependent on features of the work, the income it generates, the nature and structure of the job, its timing and total hours, and on the environments and relationships that children experience when they are not in the care of their mother i.e. through child care.

3.3.1 Summary

Although the review conducted by Shonkoff and Phillips highlights several important variables that may act as confounding variables in the current model, the review was published in 2000 and thus may miss more recent gains in knowledge. As a result, in this chapter the review is updated to build in any new evidence or additional factors important for child outcomes which could contaminate the model. Because of the wealth and diversity of information surrounding this topic it would be impossible to

summarise everything in a systematic way. Instead remainder of this chapter forms a 'review of reviews', discussing the findings of reviews that have explored development in the early years, either related to interventions or on an observational basis.

3.4 Review of reviews

3.4.1 Aim

The aim is to identify relevant potential confounders which could independently affect child cognitive and behavioural outcomes. Controlling for confounders allows greater confidence in the accuracy of the model outcomes, meaning that any effects observed relating to maternal postnatal depression are more likely to be true effects.

3.4.2 Method

3.4.2.1 Search strategy

Several approaches were used to locate studies for inclusion in the review. The principal method of literature identification involved a search of online databases, including PsycINFO, Medline, Embase, and Cochrane, covering reviews published from 2000 to June 2013. All combinations of key words in the following groupings were used: (child, childhood, infant, toddler, early years), (outcome, outcomes, development), and (review). Search tools such as medical subject headings (MeSH), Boolean operators and truncation symbols were used to expand and narrow searches. An example search strategy appears in appendix 2. Initially the search mapped key words to abstracts, titles and study key words, however due to the broad scope of the research question this yielded a result which would be unfeasible to analyse within the scope of this thesis. Consequently search terms were restricted to titles only to reduce the number of results to a more practical number. As a result of this methodological limitation it is possible that relevant reviews may have been missed, however the scope of the search makes it likely that the reviews included in the present review are still at least a representative sample of the total body of potentially available research. In addition to database searching, various combinations of key words were entered into Google Scholar to uncover any additional publications or grey literature. Because this chapter takes the

form of a review of reviews, it was not necessary to hand search bibliographies of relevant articles as they would unlikely document additional review material. It was the intent to locate any original reviews where a review was noted as being an update; however no reviews of this nature were identified.
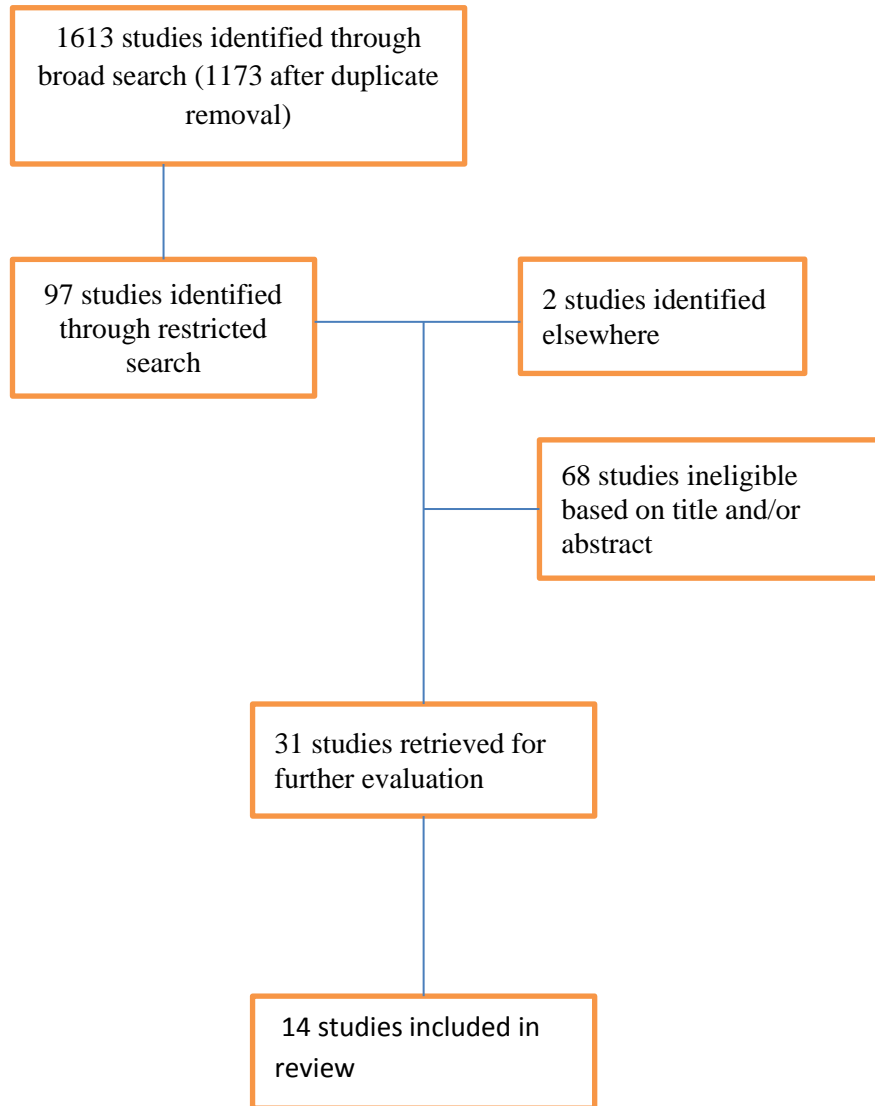
Figure 3.1. Flow diagram of inclusion/exclusion of literature process

3.4.2.2 Inclusion/exclusion criteria

To be included in the present review, reviews had to meet the following criteria.

- Focus on child outcomes and early years circumstances
- Be published from the year 2000 onwards to ensure recent and relevant data

- Have empirical methodology and narrative, systematic, or meta-analytic syntheses (theoretical or narrative literature reviews without a search strategy were excluded)
- Describe outcomes of normally developing children without physical or severe mental impairment.

3.4.2.3 Quality assessment

The quality of identified reviews was assessed using the AMSTAR (Assessment of Multiple Systematic Reviews) tool for assessing systematic reviews. This is a 11-item assessment tool developed by Shea and colleagues (2007) by combining questions from the enhanced Overview Quality Assessment Questionnaire (OQAQ) with a checklist created by Sacks (Sacks et al., 1987) in addition to three further novel items. Questions are directed towards important facets of review design such as search strategy and quality assessment of included studies. The tool has been shown to have good agreement, reliability, construct validity, and feasibility (Shea et al., 2007, 2009).

Although many of the included reviews are not specifically labeled 'systematic', the key quality items are still relevant and are able to provide a good indication of how well each review has been carried out. A copy of the AMSTAR screening tool is included in appendix 2.

3.4.2.4 Information extraction

Standard information was extracted from each review, this included;
- Age of child at outcome measurement (exposure measurement where applicable)
- The aim of the review
- The type of child outcome assessed
- Important covariates for child outcomes
- Key findings

In order to facilitate quality assessment of reviews additional information was extracted, such as whether a quality assessment of included studies was carried out and whether a comprehensive search strategy was undertaken.

3.4.3 Results

The included reviews are summarised in Table 3.1. For a complete table inclusive of study findings please see appendix 2.

Table 3.1. An overview of the included studies including quality assessment rating

Author	Theme	Number of studies	Outcome measure	Important covariates for the model	Quality
Breinholst et al. (2012)	CBT for child disorders, why parental involvement has not enhanced outcomes	N=11 Type=RCTs	Outcomes of child CBT according to parental involvement	Gender, age, parental mental health, parenting style (attachment/warmth)	Moderate
Peacock et al. (2013)	Effectiveness of home visiting on child outcomes	N=21 Type=unclear Quality criteria seem to relate to RCTs	Child abuse and neglect (n=6) developmental delays (n=11) health (n=10)	Parity, maternal age, child age, maternal depression, birth weight (important for future health), socio-economic (intervention target)	High
Orton et al. (2009)	Early intervention s for preterm infants	N = 27 Type: any	Cognitive and motor outcomes	No covariates mentioned	Weak
Ali (2012)	Relationships with parents and teachers and outcomes	N= 17 Type=any quantitative study (qual excluded)	Academic achievement – unknown how this is assessed. Behaviour – unknown how this is assessed.	Parental-child relationship (attachment), teacher-child relationship, child gender, ethnicity, birth order, SES, parent occupation, education level, parent age	Moderate
Deault (2010)	Parenting in relation to ADHD	N=22 Type=correlation, observational, longitudinal	Behaviour and academic outcomes, co-morbid problems with ADHD	Parenting (warmth), maternal mental health, family conflict	High

Table 3.1. An overview of the included studies including quality assessment rating ctd.

D'Onise et al. (2010)	Preschool and child health outcomes	N=37 Type=RCTs, quasi-experimental and retrospective cohort.	Child health – disease, disease risk factors, health behaviours, indicators of wellbeing.	Income below poverty line, low maternal IQ (previously associated with poor child outcomes).	High
Dewalt and Hink (2009)	Health literacy and child outcomes	N=24 Type=any	Health care use, general health outcomes.	Ethnicity, maternal depression, socio-economic status	High
Jelleyman and Spencer (2008)	Residential mobility	N=22 Type=any	Cognition, behaviour, mental health, general health, service use	Ethnicity, age, gender, parental mental health, single parenting, socio-economic position, neighbourhood	High
Poobalan et al. (2007)	Treating maternal depression and mother-infant interaction	N=8 Type=RCTs, controlled clinical trials	Parent-infant interaction, cognitive and behavioural development	Maternal education, marital status, employment status, comorbid maternal mental health problems	High
Thomas and Zimmer-Gembeck (2007)	Behavioural outcomes of parent-child interaction therapy	N=24 Type=cohorts	Parenting problems/stress, Child behaviour problems	No covariates – highlights importance of parent-child interaction	Moderate
Chaimay et al. (2006)	Risk factors associated with child language problems	N=15 Type=any (case control, cross-sectional, longitudinal	Language development – specific language impairments and delays	Antenatal care, birth weight, breast feeding, gender, birth order, parental education, single parenting, poverty, ethnicity where multiple languages used	Moderate

Table 3.1. An overview of the included studies including quality assessment rating ctd.

| Anderson et al. (2003) | Effectiveness of early childhood developmental programs | N=16

Type=not clear, all interventions | Cognitive, behaviour, risk behaviours, general health | Poverty – target of intervention, no actual covariates listed. | Moderate |
|---|---|---|---|---|---|
| Grace et al. (2003) | Maternal mental health and child outcomes | N=13

Type=any – all longitudinal or prospective | Cognitive and behavioural development | Child gender, family socio-economic status, maternal education, maternal age | Moderate |
| Goodman et al. (2011) | Maternal depression and child psychopathology | N=93

Type=any quantitative | Child behaviour; internalising, externalising, general, negative effect/behaviour, positive effect/behaviour | Child age, gender, ethnicity, income/poverty status, maternal age, single parenting | Moderate |

Table 3.2. A summary of important covariates for the model identified in the reviews

Important contexts for child development	Evidence
Family resources/demographics	
Parental employment	Poobalan et al., (2007)
Poverty status	Peacock et al., (2012) Ali, (2012) D'Onise et al., (2010) DeWalt & Hink, (2009) Anderson et al., (2003) Chaimay et al., (2006) Jelleyman et al., (2007) Grace et al., (2003) Goodman et al., (2011)
Parental Education	Ali, (2012) D'Onise et al., (2010) Chaimay et al., (2006) Poobalan et al., (2007) Grace et al., (2003)
Family structure (Single vs. two parents)	Chaimay et al., (2006) Jelleyman et al., (2007) Goodman et al., (2011)
Neighbourhood quality	Jelleyman et al., (2007)
Maternal age at child birth	Peacock et al., (2013) Ali, (2012) Goodman et al., (2011) Grace et al., (2003)
Ethnicity	Ali, (2012) DeWalt & Hink, (2009) Jelleyman et al., (2007) Goodman et al., (2010)
Parity	Peacock et al., (2013) Ali, (2012) Chaimay et al., (2006)
Child characteristics	
Child gender	Breinholst et al., (2012) Ali, (2012) Chaimay et al., (2006) Jelleyman et al., (2009) Grace et al., (2003) Goodman et al., (2011)
Child age	Jelleyman et al., (2007) Peacock et al., (2013) Breinholst et al., (2012) Goodman et al., (2011)
Birth weight	Chaimay et al., (2006)
Breast feeding	Chaimay et al., (2006)

3.4.3.1 Quality assessment

Reviews were generally deemed to be of good quality, with all reviews performing a comprehensive search strategy and data extraction, as well as providing details of included studies. However, five reviews did fail to state whether included studies were quality assessed reducing the level of quality allocated to the review, these were allocated a 'moderate' rating. A direct split was found with half of the reviews scoring 'high quality' and the other half scoring 'moderate quality'.

3.4.3.2 Family characteristics

Poverty

Nine of the 14 reviews included in the present review cite poverty as an important factor for child development (Peacock et al., 2012, Ali, 2011, D'Onise et al., 2010, DeWalt and Hink, 2009, Anderson et al., 2003, Chaimay et al., 2006, Jelleyman et al., 2007, Grace et al., 2003, Goodman et al., 2011). Three of these reviews explored interventions targeted at families with low socio-economic status (Peacock et al., 2013, D'Onise et al., 2010, Anderson et al., 2003) all of which showed that receiving some form of intervention was beneficial to child outcomes. Peacock et al., (2013) examined the effectiveness of home visiting interventions by nurses, midwives or other health professionals, in a set of 21 studies with varying samples in developed countries. Although the included studies found that home visiting interventions were associated with improved parenting, improved cognition, language and general health development in young children, these differences were most often insignificant, suggesting a pervasive effect of poverty resistant to this type of intervention.

In contrast to this D'Onise et al., (2010) found much greater beneficial effects of early interventions where the target of the majority of intervention groups were families with low income or below poverty line (76% of included studies). In this case preschool based interventions were found to be effective at mitigating effects of poverty, showing beneficial effects of preschool interventions on child mental health, social competency and future crime prevention. In this review studies had a much longer follow up time (up to 21

years) allowing for more longitudinal outcomes to be recorded. Such results were also replicated in an earlier review by Anderson et al. (2003) also exploring the effects of preschool programs on child outcomes. Although the interventions examined between reviews differ, there may be additional effects of timing in terms of child age that are important for analysing child outcomes, affirming the importance of assessing multiple covariates when investigating child outcomes.

Chaimay et al., (2006) used reviewing methodology to explore the risk factors associated with child language problems. Although parental education and child factors such as gender were found to be of greater importance, family characteristics such as living in poverty were highlighted as modest risk factors of child language problems.

The remaining 6 reviews did not explore poverty or socioeconomic status directly but cited it as an important covariate (Ali, 2011, DeWalt and Hink, 2009, Chaimay et al., 2006, Jelleyman et al., 2007, Grace et al., 2003, Goodman et al., 2011), moderating the effects on child outcomes of; parent-child relationships (Ali, 2011), maternal depression (Grace et al., 2003, Goodman et al., 2011), parent literacy (DeWalt and Hink, 2009), and residential mobility (Jelleyman et al., 2008).

Although it is not necessarily being raised in poverty that causes these effects, but may be other variables correlated with poverty such as parental education and employment, it still appears to be a highly important variable that could influence child outcomes and thus should be included in models exploring the effects of exposures, such as maternal depression, on child outcomes.

Maternal education

Five reviews referred to maternal education (Ali, 2012, D'Onise et al., 2010, Chaimay et al., 2006, Poobalan et al., 2007, Grace et al., 2003). Only one study shows the direct effects maternal education can have on children (Chaimay et al.) whilst the others list it as a potentially important covariate. Parental education can have direct consequences for child outcomes, whereby children born to parents with low educational levels are around 2-3 times more likely to have language impairment than those who were from parents with a higher degree (Chaimay et al., 2006). These findings were based on 3 studies, two of

which had rather small sample sizes (<200) but the third was based on a large cohort of almost 6000 children allowing for greater faith in the results.

The importance of maternal education is represented in the other reviews in different ways. Firstly D'Onise et al. explore the effects of preschool based interventions on child outcomes. These interventions were specifically targeted at families characterised as in poverty with low levels of maternal education. This sample was targeted as a result of much literature indicating this to be a high risk group for poor child outcomes.

Two further reviews exploring the relationship between maternal depression and child outcomes both suggest that maternal education can mediate the effect depression has on child outcomes (Poobalan et al., 2007, Grace et al., 2003) by explaining some of the variance. This may be through direct effects as asserted in the review by Ali (2011) which suggests that maternal education may have effects on children's academic achievement. The association between parental education and children's achievement and behaviour has been cited as the most replicated finding from developmental studies (Shonkoff and Phillips, 2000). Low maternal education is also often found to be co-occurring with other negative contexts such as poverty, thus suggesting it may interact with other negative circumstances and be a part of cumulative circumstances that affect child outcomes.

Family structure

Three reviews noted single-parenting as an important factor for child outcomes (Jelleyman et al., 2007, Chaimay et al., 2006, Goodman et al., 2011). Each review explored a different theme. Jelleyman and colleagues explored the effects of residential mobility on child cognitive, behavioural and general wellbeing. Results from 22 studies showed that residential mobility interacts at neighbourhood, family and individual levels, with higher rates of residential mobility associated with poverty, parental unemployment, family disruption and single parenting. These factors were shown to have a cumulative effect on child outcomes, with children exposed to all of these issues faring worse. These findings are somewhat duplicated in the work of Goodman and colleagues where the focus is on maternal depression rather than residential mobility. Again a cumulative effect of circumstances was found with samples that included more single-parent households yielding higher effect sizes for the association between maternal depression and children's

externalising and negative behaviours. In contrast to these reviews, Chaimay et al. (2006) show that children born from unmarried mothers are 1.5 times more likely to have a language impairment than children born from married mothers suggesting that although being a single parent can enhance the effects of other childhood circumstances, it can also act in isolation to evoke negative outcomes.

Employment

Only one review cited maternal employment as an important variable for studying child outcomes (Poobalan et al., 2007). In this review the effects of various interventions such as counselling targeted at improving the maternal-child interaction in view of improving child outcomes were assessed. Mothers' work status was cited as being a potential confounder in explaining differences in mother-child interaction outcomes between treated and untreated groups, however no explanation of why this may be the case was offered. One possible explanation may be that children of mothers who return to work soon after birth may be at increased risk of failing to form a secure attachment due to the decreased period of time spent with the mother. As noted above, it may also be that mothers returning to work early are more likely to become depressed as a consequence of separation from their infant, further hindering the mother-child relationship and parenting style. Employment is thus an important variable to be considered for the model as it could contribute both to the development of depression and independently to the child's development via impairments in parent-infant interactions.

Maternal age at child birth

Two reviews cite maternal age at birth as being important for child outcomes (Goodman et al., 2011, Grace et al., 2003), both of which were focused on the effects of maternal depression on child outcomes. Goodman and colleagues found women who give birth as teenagers have higher rates of depression symptoms, poverty, and were more likely to be single parents compared to older mothers, again lending to the idea of cumulative exposures. Due to a low number of studies reporting maternal age they were unable to effectively examine the effect of maternal age on child outcomes, however given the factors associated with being a young mother it is likely to be an important covariate for building models exploring child outcomes. Similar to Goodman et al., the review by Grace

and colleagues also found that few studies report maternal age despite it being a potentially important covariate with only 1 included study incorporating maternal age as a covariate when modelling child outcomes. Although this provides limited evidence for the importance of maternal age, the fact that it is associated with several other covariates that have been highlighted as important for child outcomes it is still likely to be an important variable for modelling child outcomes. This could be considered in line with the accumulation model described in Chapter 2 which identified that risk exposures may be clustered (Chapter 2, Figure 2.2).

Ethnicity

Four reviews cited ethnicity as an important covariate (Ali, 2011, DeWalt and Hink, 2009, Jelleyman et al., 2008, Goodman et al., 2011). Despite identifying ethnicity as a potentially important confounder only one review attempts to explain why this might be (Goodman et al., 2011). It is suggested that ethnic group should be considered a demographic variable that can contribute to the definition of the context of the lives of children and can increase the probability of maternal depression. In addition to being related to maternal mental health, being from an ethnic minority is also often associated with a range of other negative circumstances such as poverty and limited access to other resources and in some countries such as the US, health care. This idea has been verified through a large literature review conducted using UK literature (Netto et al., 2011). This suggests that ethnicity could mediate other factors that may pose a cumulative risk to child outcomes. Another possible explanation that is not explored by Goodman et al. is the notion that there may be cultural differences in parenting that may affect child outcomes. Research carried out by Ho et al. (2008) show ethnicity related differences in the association between parenting and child outcomes in that parental harshness was positively related to child aggression in white Canadian families but negatively related in Asian Canadian families.

Parity/birth order

Although none of the included studies examined parity or birth order directly, three reviews cited it as a potential confounder (Peacock et al., 2013, Ali, 2011, Chaimay et al., 2006). One study did not discuss parity or birth order in detail but noted that it was often cited as a possible confounder in the study of parent-child relationships (Ali , 2011). In

another review older mothers with previous children were identified as vulnerable to depression (Peacock et al., 2013) suggesting parity may have an indirect effect on child outcomes through maternal mental health. Alternatively research has shown that subsequent children often receive less quality time with parents than first children which may impact on attachment (Price, 2008). As discussed in Chapter 2, attachment is highly important for development, reiterating that parity may indirectly affect child outcomes.

The final review commented directly on the effects of birth order on child outcomes (Chaimay et al.). In this review two studies were found, both suggesting that birth order can have direct effects on children's language development (Horwitz et al., 2003, Stanton-Chapman et al., 2002). In the first study, first born children were noted to be two times more likely to gain benefits for language developmental skills than later children (Horwitz et al., 2003). Building on this the second study showed that later born children are 1.5 times more likely to have a specific language impairment than first born children (Stanton-Chapman et al., 2002).

Taken together these studies show the direct and indirect effects of parity/birth order, ascertaining it to be an important covariate in child outcome study.

3.4.3.3 Child characteristics

Child gender

Six reviews documented potential differences in outcomes attributed to child gender (Breinholst et al., 2012, Ali, 2011, Chaimay et al., 2006, Jelleyman et al., 2009, Grace et al., 2003, Goodman et al., 2011).

In one review gender differences were found in the relationship between parenting style and child academic achievement (Ali, 2011) though which gender was most affected varied between included studies suggesting conflict within the literature. Such differences were also found between parenting style and levels of child anxiety (Breinholst et al., 2012) with girls being more negatively affected than boys. In their study of the effects of neighbourhoods on child outcomes, Jelleyman and colleagues also noted that there may be differential effects according to gender, again with boys been most affected.

Gender differences in child outcomes were also noted in relation to maternal depression (Grace et al., 2003, Goodman et al., 2011). Grace and colleagues found that maternal postnatal depression holds a significant impact for boys' cognitive development, as measured by the Bayley scales, which was not found for girls. This may be due to inherent gender differences in development or may be due to differential parenting. One of the interesting findings arising from the review by Grace et al. was a significant difference in maternal speech related to the gender of their child. For depressed mothers with male infants maternal speech was found to be significantly less infant-focused than depressed mothers with female infants. This again goes back to the importance of parenting and the relationship between mother and child.

In contrast to this, Goodman and colleagues comment on the effects of maternal depression on behavioural outcomes in terms of child gender, noting somewhat conflicting evidence. For boys there is some limited evidence of increased risk for externalising behaviour problems in relation to maternal depression compared to girls, suggesting boys may be more negatively affected. In contrast to this there are much more consistent findings suggesting that female children of depressed mothers are at greater risk than males for later depressive symptoms. It appears from this review that there may be some differences in childhood outcomes according to gender but the nature of these is not clear. As suggested by Grace et al., differential effects may be a result of differences in parenting carried out by the mother.

As well as interacting with other factors such as maternal depression, gender differences were also noted in general child development. One review shows outcomes directly linked to gender (Chaimay et al., 2006). In this review exploring risk factors associated with language development, specific language impairments were found to be four times more prevalent in males than females. Boys were also noted to be more sensitive to maternal language impairments where mothers with specific language impairments had seven times as many boys and four times as many girls with specific language impairments.

Although the research is conflicted, and there is no theoretical explanations of why gender differences may exist there does appear to be some effect of gender for certain outcomes. Taking gender into consideration in statistical analysis may be important and failing to do so may mask important effects on child outcomes if they are in fact gender specific.

Child age

Four reviews noted differences in outcomes according to the age that children were exposed to negative circumstances (Jelleyman et al., 2008, Peacock et al., 2012, Breinholst et al., 2012, Goodman et al., 2011). In Jelleyman and colleagues' review relating to the outcomes associated with residential mobility, age was found to be a critical factor for mediating outcomes. When residential mobility took place at a younger age, at infancy or pre-school, no long-term effects were seen for development. However, once children reached school age and adolescence, significant effects were observed linking higher levels of residential mobility with behavioural problems. These findings suggest that there may be a critical period where residential mobility is important conforming to the critical periods model described in Chapter 2.

Similar effects were found with regards to the timing of interventions. In one review exploring the outcomes associated with home visiting interventions, better effects on developmental outcomes were observed when children were exposed to the intervention at a younger age. This may relate back to the idea of critical periods for development discussed in Chapter 2, suggesting that the developmental stage of the child is important for determining exposure effects.

In another review examining the effect of parenting and parental involvement in child treatment for child disorders, age is again shown to be important (Breinholst et al., 2012). In this review the impact of parental involvement is dependent on age, with results showing that younger children who are more dependent on their parents and spend more time with them achieve better results from intervention with parental involvement compared to older children.

The final review comments on how the timing of exposure to maternal depression affects child outcomes (Goodman et al., 2011). Findings from the included studies showed that the younger the children's age at first exposure to their mothers' depression will have a stronger negative impact than later first exposure. Consistent with this finding, effect sizes for the association between maternal depression and children's internalizing and externalizing problems were negatively correlated with children's age, thus it appears that

early onset of maternal depression holds greater meaning for child outcomes than later onset, again emphasising the importance of age of exposure to negative circumstances.

Collectively, all reviews referring to the age of the child document stronger negative effects associated with earlier exposure providing strong evidence of the early years being a critical period for development and a period where children are highly sensitive to negative exposures.

Other variables

In addition to the variables listed above several other factors that may influence child outcomes were identified.

One review listed birth weight as being a potential predictor of language development (Chaimay et al., 2006). In this review four studies reported on this association, three of which suggested that birth weight may enhance the risk of language impairments whilst the final study cited inconclusive evidence. Where differences were found these were often small. As well as low birth weight, whether children were breast fed was also identified as a potential predictor of language development in the same review. Only one study was used to support this assertion which showed that breastfeeding can reduce the risk of language impairments but this was not significant.

3.4.4 Summary

Findings from the review of reviews identified several important variables which could both independently or collectively influence child outcomes and thus need to be included in the model assessing the impact of maternal mental health on child development (Figure 3.2).
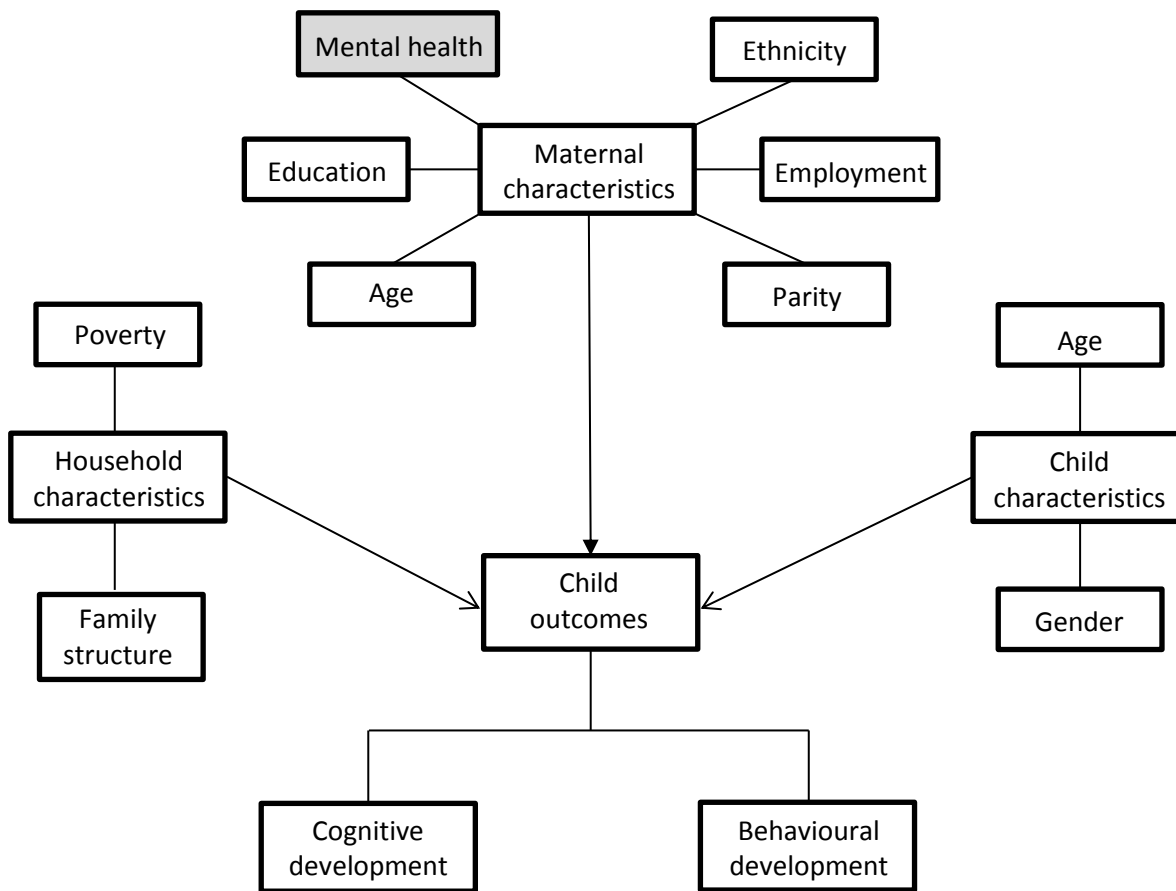
Figure 3.2. Important covariates for child outcomes to include in statistical modelling

As can be seen in Figure 3.2 both family level and child level characteristics were identified. The results of the review highlighted the maintained importance of those variables identified by Shonkoff and Phillips (2000) with current literature still emphasising these variables as important for child outcomes. In addition to this several new factors were also identified. These were additional maternal characteristics: ethnicity, age at child birth, and parity, as well as child characteristics: age and gender.

As discussed in each of the corresponding sections above, most of the factors identified appear to have an independent contributory effect to child development. However, many of these are likely to be co-occurring and may be markers for another factor that has a more direct causal influence on development. For instance, the literature showed an association between maternal age and child behaviour problems. It is unlikely that being born of a younger mother has a direct causal influence on behavioural development.

Rather, young mothers are more likely to be less educated, and are more likely to be living in poverty. Both of which have more direct links with child outcomes, as described above. Ethnicity is another of these factors. Again it is unlikely that ethnicity directly affects development but is a marker for other risk factors such as poverty. Together these factors may have a cumulative effect on child development.

In addition to being risk factors for child developmental problems, a number of the factors may also contribute to the development of maternal depression, acting as stressors (as discussed in Chapter 1). For instance, being a younger mother is linked to maternal depression. This is likely to be a marker for poverty and being a single parent, both of which are thought to be potential stressors contributing to the development of depression (Kendall-Tackett, 2009).

As it is difficult to distinguish which factors are likely to have a direct causal association with child development, and which are markers of other risk factors, all identified factors will be included in the modelling throughout this thesis as covariates.

Chapter 4: The Importance of Maternal Mental Health Diagnosis for Child Developmental Outcomes

4.1 Introduction

In this chapter, a recent birth cohort of UK children is used to examine how maternal mental health affects the behavioural development of offspring up to the age of seven years old. The focus of this chapter is assessing the effect of a missed diagnosis of maternal depression during the postnatal year on children's cognitive and behavioural development over time. This effect is compared with that of children whose mothers were diagnosed and treated for postnatal depression, as well as both groups being compared with children of mothers who showed no indication of depression during the first postnatal year. The primary outcome measures for this chapter are child behavioural development as measured using the 'Strengths and Difficulties Questionnaire' (Goodman, 2001) and child cognitive development as measured by the vocabulary subscales from the British Ability Scales (Elliot, 1983).

4.1.1 Background

As discussed in Chapter 1, maternal mental disorders such as depression occurring throughout pregnancy and the first postnatal year can have serious consequences for individual women, their partners, babies and any other children for whom the woman is responsible (Redshaw and Van den Akker, 2007). Maternal postnatal depression is recognised as a major public health issue because of its globally high prevalence (Almond, 2009) with high rates occurring amongst women in both developed and developing countries (Halbreich and Karkun, 2006).

Although carried out almost 20 years ago, the most predominantly used estimate of prevalence for postnatal depression in the UK is that generated through a large scale meta-analysis conducted by O'Hara and Swain (1996) which suggests that 14% of all mothers are affected. Considering that in the UK more than 700,000 women deliver per year

(Office for National Statistics, 2011) this estimate would suggest that around 98,000 women will fall into this category per year.

However, as described in Chapter 1 there is emerging evidence to suggest that postnatal depression is an under-diagnosed condition, thus the actual rates may be much higher (Hearn et al., 1998, 4Children report, 2011).

Given the negative consequences postnatal depression can have on children as well as for the mothers themselves, this is of great concern as effective identification and treatment of depressed mothers may prevent these negative effects.

This chapter will assess the differential effects of undiagnosed and diagnosed maternal depression on child language development as an indicator of cognitive development using the language subscales of the British Ability Scales (as discussed in Chapter 3) and on behavioural development used the Strengths and Difficulties Questionnaire (as discussed in Chapter 3).

4.1.2 Aims

This analysis aims to assess the impact of undiagnosed and consequently untreated maternal depression on child cognitive and behavioural outcomes up to the age of seven using data from the Millennium Cohort Study, and to compare this to the effect found as a consequence of diagnosed and treated depression. These two groups are compared to children of mothers not reporting depression at nine months after birth to identify deviances from normal development.

4.1.3 Objectives

Primary

- To explore the effect on child behavioural and cognitive development over time of undiagnosed postnatal depression.

Secondary

- To investigate the impact of undiagnosed postnatal depression on individual facets of behaviour, namely; externalising and internalising behaviour types.
- To identify whether child gender mediates the impact of maternal depression.

4.1.4 Research Questions

- What is the difference in behaviour according to the SDQ of children (up to age 7) whose mothers have been formally diagnosed with postnatal depression compared to those whose mothers exhibited symptoms of depression but were not diagnosed and receiving treatment?
- Is there a gender difference in the behavioural outcomes of the children related to maternal depression?
- How does maternal depression affect children's scores on the internalising and externalising subscales of the SDQ? Are there gender differences between subscales?

4.2 Methods

4.2.1 Data; Characteristics of the MCS data

This investigation uses secondary data extracted from all four data sweeps conducted to date of the Millennium Cohort Study (MCS) described in Chapter 3. Data sweeps were carried out when the child was around 9 months, 3 years, 5 years and 7 years with the majority of children at sweep 4 being around 6.3 years to 8.2 years at interview. By sweep 4, 61% of the original sample remained active in the study (n=11, 721), a moderate reduction from the first sweep of 18, 552 families (96% of the original sample). The present study uses a subset of the remaining cohort to include only families whereby the natural mother of the cohort child was the main respondent and remained active in the study at all surveys to date (MCS 1-4). This resulted in a sample size for analysis of 11,162. Where multiple birth children were included in the cohort only one child per family was included in the analysis to avoid repeated testing of the same mother. A large

number of baseline characteristics were gathered at the initial sweep and were updated at further sweeps to account for changes over time. Variables identified in Chapter 3 as important were included in the analysis and are described in Table 4.1. The average age of mothers at the cohort member's birth was 29.1 years, however there was a large range of 38 years between the youngest mother (13 years old) and the oldest mother (51 years old). As the MCS specifically targeted ethnic minorities these are well represented in the data with 11.5% of the population being non-white. In addition to oversampling from ethnic minorities, mothers from a low-income household are also adequately represented as almost a third of the sample population is classified as being in poverty according to median income levels.

Table 4.1. Baseline characteristics at cohort member age 9 months (Sweep 1)

	N	Values (%)
Total observations	**11162**	
Maternal characteristics		
Age at birth	**11,158**	Mean: 29.1 (SD: 5.7 years)(Min: 13; Max: 51)
Qualifications	**11,148**	
NVQ Level 1	867	7.78
NVQ Level 2	3245	29.11
NVQ Level 3	1675	15.03
NVQ Level 4	3403	30.53
NVQ Level 5	424	3.80
Overseas only	248	2.22
None of these	1286	11.54
Lone parent	**11,162**	
Yes	1509	13.52
No	9653	86.48
First live birth	**11,063**	
Yes	4590	41.49
No	6473	58.51
In employment	**11,157**	
Yes	5633	50.49
No	5524	49.51
Ethnicity	**11,136**	
White	9855	88.50
Mixed	82	0.74
Indian	252	2.26
Pakistani	519	4.66
Black	294	2.64
Other ethnic group	134	1.20
Poverty	**10,339**	
Not in poverty	7,373	71.31
In poverty	2,966	28.69
Child characteristics		
Child age (months)	**11,162**	Mean: 9.2 (SD: 0.5 months)(Min: 8; Max: 12)
Gender	**11,162**	
Male	5645	50.57
Female	5517	49.43
Ethnicity	**11,136**	
White	9709	87.19
Mixed	278	2.50
Indian	250	2.24
Pakistani	517	4.64
Black	284	2.55
Other ethnic group	98	0.88

Note: Age at birth refers to the mother's mean age at the cohort members' birth.

4.2.2 Weights, response and attrition

The most important design feature of the MCS to note is the sampling design. A strategic sampling design was employed which oversampled for ethnic minorites and families with a lower income meaning that survey weights should be used in any analysis. Survey weights are required when analysing survey data which is not self-weighting. Each participant is assigned a value that indicates how much each case will count for in a statistical procedure. For example, a weight of '2' means that the case counts in the dataset as two identical cases. The weight component therefore provides information about how many members within the sampling unit are represented by each member selected in the sample. Weights available in the MCS dataset are longitudinal and include both sampling adjustment and attrition/ non-response adjustment for all four sweeps. Research utilising the MCS thus far has indicated several variables related to attrition/non-response including being non-white, being a lone parent, having fewer educational qualifications, having low income, and being a young parent (Plewis and Kentende, 2006). It is important to consider the non-response of families with these characteristics as it is possible they may also serve as indicators of mothers with postnatal depression that could potentially be missing from the present analysis.

4.2.3 Variables

The independent variable of this study is maternal depression/distress. The primary outcome variable is a standardised measure of behaviour problems in children (Strengths and Difficulties Questionnaire (SDQ)). Cognitive development is also measured using the British Ability Scales (BAS). Several potential confounder variables are also included. All variables are described in more detail below.

4.2.3.1 Maternal depression

Information on maternal depression was drawn from the first survey in order to investigate whether a difference emerges in the behavioural development of children of mothers with diagnosed depression or depressive symptoms without diagnosis on a self-reported scale reported during the postnatal year.

Cases of diagnosed depression were identified from information gathered during the parental interview. Respondents were asked 'Has a doctor ever told you that you suffer

from depression or serious anxiety?'. Following a positive response respondents were then asked 'And are you currently being treated for this?'. Current cases of depression at the time of survey are therefore determined by an affirmative response to both questions.

In order to determine mothers who were not diagnosed by a doctor but who were experiencing depressive symptoms, information from a widely used mental health measure, the Malaise Inventory (Rutter, 1970), was extracted.

Box 4.1. Malaise Inventory items

Malaise Inventory

Do you feel tired most of the time?
Do you often feel miserable or depressed?
Do you often get worried about things?
Do you often get into a violent rage?
Do you often suddenly become scared for no good reason?
Are you easily upset or irritated?
Are you constantly keyed up and jittery?
Does every little thing get on your nerves and wear you out?
Does your heart often race like mad?

The Malaise Inventory is a 9 item self-completed scale whereby high levels of psychological distress have previously been categorised by a score of four or more points (Bartley et al., 2004, Mensah and Kiernan, 2010). Questions such as *'do you often get worried about things?'* and *'do you often feel depressed?'* are presented in a yes or no format (see Box 4.1). Traditionally the Malaise Inventory consists of a 24 item scale which has been broadly used as a measure of distress and depression levels (Grant et al., 1990, McGee et al., 1986) and has widely been shown to have good reliability and validity (Grant et al., 1990). The Malaise Inventory has been used in both large scale surveys of samples drawn from general populations (Dodge and Silva, 1980) and in smaller more constrained sub-groups, e.g. mothers of severely disabled children (Hirst and Bradshaw, 1983). Although the full version of the Malaise Inventory has been rigorously tested, the MCS employs a reduced 9 item version of the scale which has not been robustly validated. However, the reduced measure has been used in other comparable literature (Bartley et al., 2004, Mensah and Kiernan, 2010, Kiernan and Huerta, 2008) and although it does not

guarantee a diagnosis of depression, it can provide a good inference of the level of symptoms a mother is experiencing. Additionally, the correlation between Malaise scores and clinical diagnosis was assessed in the current sample and a moderate correlation was found (r=0.3650, p<.0001) indicating that the reduced Malaise Inventory used in this research is likely to be a good predictor of depression. However, it is important to note that it is not the intention to derive a clinical diagnosis of mental illness from this self-report data, rather the intended analysis aims to recognise mental health symptoms which can be indicative of underlying mental illness as experienced by mothers who have specifically not received an official diagnosis of depression from a clinician. Throughout this thesis, women scoring above threshold on the Malaise Inventory will be deemed 'depressed'.

4.2.3.2 Behavioural development

As discussed in section 3.2.3.2, behavioural development of children was assessed at ages three, five and seven using the Strengths and Difficulties Questionnaire (SDQ) (Goodman, 1997), which comprises 25 questions addressing five dimensions of behaviour; conduct, attention/hyperactivity, emotional symptoms, peer related problems and prosocial behaviour. Each item was rated by the mother using a scale from 0 to 2 whereby if the item is somewhat true it is always scored as 1, but whether not true and certainly true are scored 0 or 2 depends on whether the item is framed as a strength or a difficulty. Responses were summed to provide a total score (max 10) for each dimension. These constructs are then separated into the externalising subscale (sum of the conduct and inattention/hyperactivity scores, max 20) and the internalising subscale (sum of the emotional and peer interactions scores, max 20). Finally a total difficulties score was derived by summing the externalising and internalising subscales (max 40). This score represents the primary outcome for the analysis.

4.2.3.3 Language development as an indicator of cognitive skills

As discussed in section 3.2.3.1, language development, as measured using subscales from the British Ability Scales (Elliot,1983), was used as an indicator of cognitive development. The measure used to indicate language development varied across sweeps to ensure age appropriate assessments. Vocabulary skills were assessed at ages three and five using the 'Naming Vocabulary' subscale, and at age seven using the 'Word Reading' subscale.

Because the scores are not directly comparable, the raw scores from each of the subscales were dichotomised in the present analysis to separate the lowest scoring 10% of children.

4.2.4 Covariates

A range of background factors are also to be included as controls in the analyses including parental/household and child attributes.

4.2.4.1 Qualifications

Information concerning any academic or vocational qualifications held by the cohort member's parents were reported at the first sweep and updated at future sweeps. Qualifications were categorised into seven bands, one accounting for the respondent holding only overseas qualifications and six levels corresponding to the UK's National Vocational Qualification scale (NVQ) ranging from no qualifications to level five representing professional or academic qualifications at degree level, nursing or other medical qualifications.

4.2.4.2 Ethnicity

Parents were asked which ethnic identity category they felt they belonged to corresponding to 6 categories used in the UK census (Office for National Statistics, 2003).

4.2.4.3 Maternal age

Parental age was given in years at the time of the cohort member's birth.

4.2.4.4 Poverty

The MCS includes income data which has been equivalised by the Organisation for Economic Co-operation and Development (OECD). Equivalisation is a standard methodology that adjusts household income to account for different demands on resources by considering household size and composition. For the purpose of this investigation income poverty will be represented by the measure that has become in the UK (DWP, 2002) and the EU (Atkinson et al. 2002) the most conventional measurement of relative

poverty; those households with net equivalent household income less than 60% of the median. This measurement has been utilised as a measure of income in work comparative to the present study (for examples see Kiernan and Mensah, 2009, Hobcraft and Kiernan, 2010). To this end income data available in the MCS includes a measurement of poverty status which identifies cohort members who were living in poverty with an indicator variable that had a value of 1 if the family's income was below the conventional poverty line (i.e. below 60 per cent of national median income before housing costs) and 0 otherwise. The dichotomous variable was preferential compared to a continuous measure in this instance as there is some evidence to suggest that the association between income and children's development is non-linear (Duncan and Brooks-Gunn 1997) and also to fall in line with the comparative work mentioned above. In the sample utilised in this analysis, 28.7 per cent of cohort members were classified as living below this poverty line in the first sweep. This figure remained relatively constant across all four sweeps, falling by just 2.3% by sweep 4 to 26.4% of cohort children living in poverty.

4.2.4.5 Employment status

The mothers' employment status at each sweep was presented as a dichotomous variable, where the mother was either in employment or not.

4.2.4.6 Lone parenting

Lone parent status was derived as a binary variable using information concerning household attributes. Mothers were deemed to be 'lone parents' if they were living alone in the home with the cohort child. This separates women who have no home support from those who have a spouse or co-habiting partner.

4.2.4.7 Parity

In the MCS mothers were not directly asked whether or not the cohort child was their first child. Fiona Mensah, a former researcher in the Department of Social Policy at the University of York developed an algorithm to estimate this by using a series of questions about previous children including; the number of siblings in the household, the number of other children not living in the household and whether the mother had ever had a stillborn

birth. This allowed for an estimate to be made as to whether the cohort child was the mother's first live baby or not (Mensah, F, personal communication, 30th November 2011). This was merged into the data by matching on MCS subject ID number.

4.2.4.8 Child ethnicity

As child ethnicity can vary from the mother this was recorded separately and was derived from the parental interview using the same 6 census categories as were used to determine the mother's ethnicity.

4.2.4.9 Child gender and age

Due to the timing of data collection for each survey there is some variance between the ages of the cohort members within sweeps. Accurate age data is of critical importance when comparing age normative data. At the 9 month survey cohort member age was collected in months, thereafter it was collected in days. For consistency purposes the age data from MCS sweeps 2-4 has been converted into an approximate age in months. The cohort member's gender was also included as a covariate.

4.2.5 Group allocation

Three groups are used in the analysis which correspond to the mother's mental health status at sweep 1 when the cohort child was around 9 months old, these are; diagnosed with depression and recieving treatment, self-assessed psychological depression without diagnosis and non-depressed.

4.2.5.1 Diagnosed and in treatment

The first group in the study comprises cases of clinically diagnosed depression. As described above, mothers who reported that they had been diagnosed with depression by a doctor and who were receiving treatment for depression at the time of the 9-month survey were classed as 'diagnosed' mothers. In following this methodology 957 women (8.57%) were categorised into this group.

4.2.5.2 Self-assessed depression without treatment

The second group in this study comprises cases that identify as experiencing depression symptoms but are not receiving treatment. Mothers were allocated to this group via one of two pathways. Firstly, mothers could report that they had never been diagnosed with depression by a doctor but score above threshold for depression on the Malaise Inventory (N=656). Alternatively, mothers reporting that they had been diagnosed with depression at some point but were not currently receiving treatment were allocated to this group if they scored above threshold on the Malaise Inventory (N=364). This gives a total of 1020 women (9.14%) in this group. Although this means that there may be some women allocated to this group who had recently been diagnosed with depression but who were not receiving treatment, the MCS does not allow for the distinction between these women due to the phrasing of the questions. It was not deemed appropriate to separate these women into two individual groups as this would result in smaller sample sizes and reduced power for the analysis. This is acknowledged in the discussion.

4.2.5.3 Non-depressed

The final group in this study acts as the control group. This group comprises women who stated that they were not currently in treatment for depression and were identified as not being depressed on the Malaise Inventory. This group comprises 8916 women (79.88%).
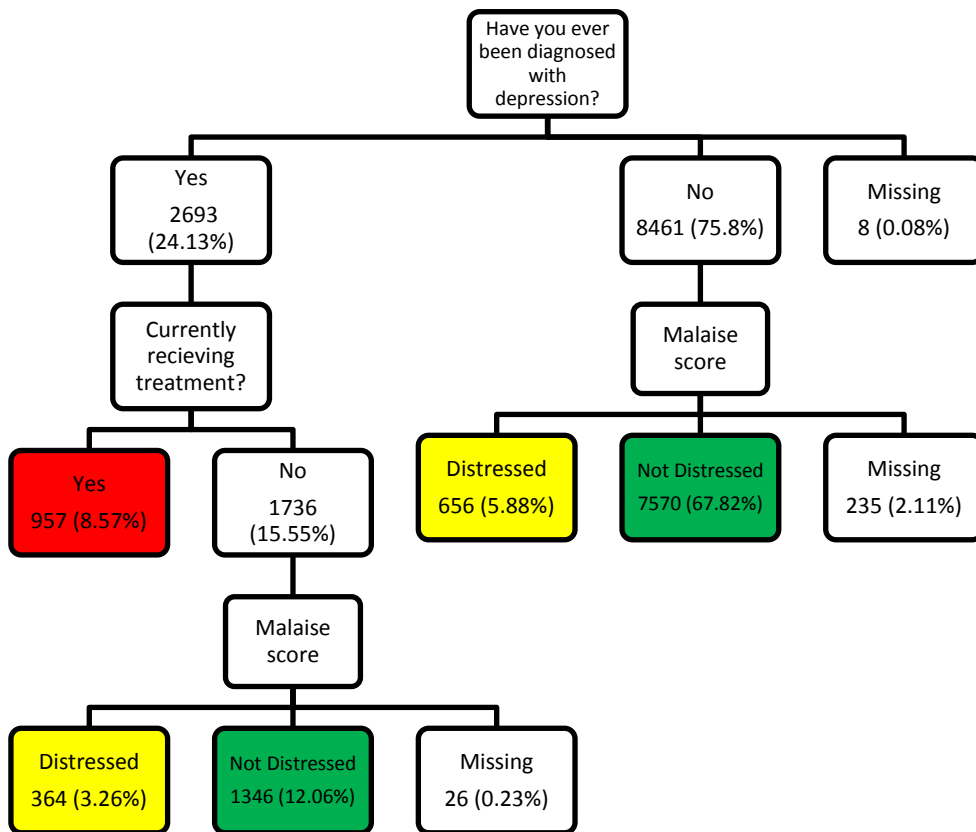
Figure 4.1. Method of study group derivation

4.2.6 Multiple imputations for missing data

4.2.6.1 Missing data

Analyses of multivariate data are frequently hampered by missing values. In the MCS data set, although 69% of the original sample remained active at the fourth sweep (Hansen et al., 2010), various within-subject data points are often missing for particular variables. For ethical reasons, participants were able to opt out of answering any interview questions they did not wish to answer. Consequently the MCS data set is subject to missing data across variables. Missing data often causes problems during analysis not only because it leads to possible bias in estimates (Newman and Sin, 2009, Gorelick, 2006), but also because the analysis loses power as subjects with missing data are often omitted from the analysis (Myers, 2000). Whilst analysing only complete cases is straightforward, the information contained in the incomplete cases is then lost. This approach further ignores possible systematic differences between the complete and incomplete cases meaning that the resulting estimates may not be applicable to the target population, especially if only a limited amount of complete cases are available. Due to the fact that the MCS sample used in this analysis is very large, it is unlikely that missing data would have a large impact on the final outcome. However, it may be the case that particular sub-groups of participants are more likely to have missing data points than others such as ethnic minorities or younger mothers, consequently if the missing data from these groups was not taken into consideration the results would not be adequately representative of the national population and thus would be less generalisable.

Using multilevel modelling offers some protection against the effects of missing data as observations at each time point influence estimates of intervention effects at other time points where an appropriate covariance pattern is specified (Brown and Prescott, 2006). Participants who thus have scores at earlier time points will thus still be taken into account for later time points. However, where there are large differences between groups further analysis may be required to account for the effects of missing data. One such solution is to apply multiple imputation techniques to generate estimates of the missing values. The approach taken depends on the pattern of missing data.

4.2.6.2 Patterns of missing data

There are several imputation methods which can be used with the method of choice being dependent on the patterns of missingness. To determine the patterns of missingness the command 'mvpatterns' can be used in STATA 11 (StataCorp., 2009). Generally data is classified as monotonic or non-monotonic/ arbitrary (Fairclough, 2010). As missing data in the MCS is generally unrelated to time and study drop-out but to particular survey questions, a non-monotonic pattern of missing outcome data was observed. For data sets with non-monotonic missing patterns, a Markov Chain Monte Carlo (MCMC) method (Schafer 1997) that assumes multivariate normality can be used to impute all missing values.

4.2.6.3 Multiple imputation

Since its introduction to survey data analysis more than thirty years ago (Rubin, 1978), multiple imputation has become an important and influential approach in the statistical analysis of incomplete data (Kenward and Carpenter, 2007). Consequently, more ad hoc practices, such as list wise deletion in which an entire record is excluded from the analysis if any single value is missing, are increasingly being recognised as insufficient measures of dealing with missing data (Allison, 2002). Multiple imputation is a simulation technique that replaces each missing data point with a set of $m > 1$ plausible values. The m versions of the complete data are analysed by standard complete-data methods, and the results are combined using simple rules to yield estimates, standard errors, confidence intervals and p-values that formally incorporate missing-data uncertainty. In the past as few as three to five imputations were deemed adequate in multiple imputation (Rubin 1996, p.480). More recent literature suggests that this number may be insufficient to generate satisfactory results (Bodner, 2008). In this instance the value of m was set as 10 in order to generate data imputed data sets. Multiple imputation was adopted for this analysis for several reasons. Firstly, it can be applied very generally, to very large datasets such as the MCS which often have complex patterns of missingness among variables. Secondly, because numerous data sets are generated and then collated the imputed data sets better account for uncertainty about the predictions of unknown values, than would other imputation measures such as single imputation. And thirdly, the imputation model is able to include variables not in the substantive model, which can lead to additional efficiency, and allows

for the inclusion of covariates which may be predictive of missingness but not important for the multilevel model. To carry out multiple imputation in this instance two procedures from the SAS package (SAS 9.2, 2010) were used, namely; the MI procedure and the MIANALYZE procedure.

Diagnostic plots were produced for each imputation to check the model fit and compare with the observed data. Mean SDQ scores and standard deviations were noted for each imputed data set to ensure they adequately represented the original data (see appendix for diagnostics and descriptive statistics for the imputed data sets). Following these checks parameter estimates were drawn by compiling the imputed data sets.
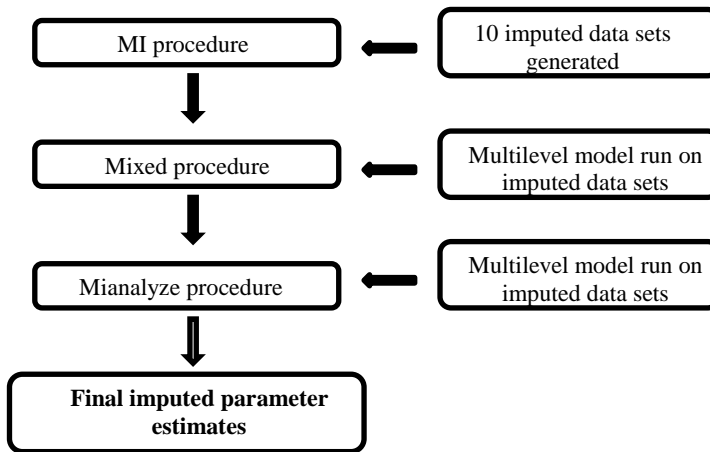
Figure 4.2. Multiple imputation analysis procedure based on SAS procedures (SAS, 2011)

4.2.7 Statistical analysis: Multilevel modelling

The present analysis aims to assess the effects of diagnosis of maternal postnatal depression on child outcomes at each of the three available time points whilst testing whether treatment interacts with time, thus an approach that deals with repeated measures is required. Fixed effects based approaches can sometimes be considered satisfactory (Brown and Prescott, 2006) whereby separate analyses at each time point are undertaken, however, this involves repeated testing and significant findings are more likely to be identified at individual time points by chance. Additionally, standard errors are also likely be less accurate as they are driven only by an individual time point rather than using data from all time points. Using a multilevel model is particularly advantageous in this respect as the model uses information from other time points in predicting estimates for another time point, thus missing data at individual time points has less of an impact on the overall outcomes of the model.

Multilevel models are appropriate for research designs where there are multiple levels or hierarchies, i.e. individuals as a lower level are usually nested within contextual/ aggregate units at a higher level. Although the lowest level of data in a multilevel model is usually an individual, multilevel models are also an appropriate method of analysing repeated measurements of the same individual. Though multilevel models can be used on data with many levels, two level models are the most common and simplistic, always examining the dependant variable as the lowest level. When measurements are repeated on the same subjects, a two-level hierarchy is established with time / repeated measures as level 1, nested within subjects (the child in the present analysis) as level 2 (Figure 4.3). As there is only one child per mother, there is no clustering of children within mothers thus 'mother' does not represent an extra level in the model. Covariates associated with the mother were treated as child level covariates, e.g. the child has a single mother or the child is being raised in poverty.
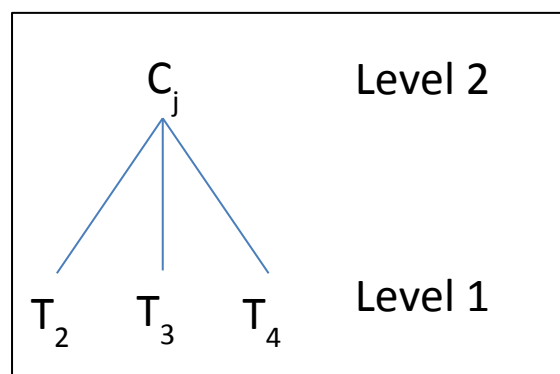
Figure 4.3. Hierarchical structure of data showing levels for MLM.

The primary benefit of adopting a multilevel approach is that overall intervention effects and intervention effects at each time point can be derived using a singular model. In multilevel models intervention effects are compared against a background of between-patient variation. Standard errors for intervention effects at individual time points are calculated using information from all time points and are therefore more robust than standard errors calculated from separate points and more power is provided to the model.

There are several types of multilevel model that can be used to fit repeated measures data, as in the present analysis, with the simplest approach being random effects modelling with subject effects fitted as random. This allows for constant correlation between all observations on the same participant. However, as correlation between observations may change over time, a covariance pattern multilevel model was deemed appropriate to allow for this.

In covariance pattern modelling the covariance structure is defined directly by specifying a covariance pattern for the model rather than by using random effects. The covariance pattern is defined within the residual matrix and is blocked by patients so that only observations on the same patient are correlated.

A number of covariance patterns exist and most are dependent on measures being taken at fixed times at even intervals. Individual covariance structures have different assumptions. For instance, in the 'general' pattern the variances of responses, differ for each time period and the covariances differ between each pair of periods. In contrast, the 'compound symmetry' model assumed all covariances are equal whereas in the 'first order auto-regressive' model the variances are equal and the covariances decrease exponentially depending on their separation. Possible covariance patterns can be found in SAS guidance documentation (Kincaid, 2005).

Instead of directly fitting covariance patterns to model, a random effects model could be used or a multivariate normal distribution could be fit to the repeated measures. A random effects model would give identical results to the compound symmetry mode and data fit with a multivariate normal distribution would give the same results as using a general pattern if data is complete. However, if there are missing data, then fitting a multivariate normal distribution is not appropriate as most statistical software ignore all information from cases with missing data (Brown and Prescott, 2006). Adopting covariance pattern based modelling overcomes this limitation by providing a flexible choice of covariance structures which can be fitted to complete or incomplete data.

Consequently, a covariance pattern multilevel modelling approach was deemed the most suitable for the present analysis. The process of statistical modelling involves three distinct stages: formulating a model, fitting the model to data, and checking the model. The MIXED procedure in SAS 9.2 was employed.

The primary assumptions underlying the analyses performed by the MIXED procedure are as follows:
- The data are normally distributed (Gaussian).
- The means (expected values) of the data are linear in terms of a certain set of parameters.

- The variances and covariances of the data are in terms of a different set of parameters, and they exhibit a structure matching one of those available.

Intervention and time interaction statements can be added to models. This results in effect sizes being calculated as the average of the intervention effects from each time point measure, in contrast to a weighted average of estimates from each time point being calculated as when these interaction terms are omitted. The decision on whether to include these is dependent on whether they are shown to be significant predictors of the model. In the present model a significant effect was observed for the interaction between sweep (time) and maternal diagnosis group (intervention) thus the interaction term was included in the model.

4.2.7.1 Selecting a covariance structure

The MIXED procedure in SAS 9.2 provides a variety of covariance structures. The most common of these structures arises from the use of random-effects parameters, which are additional unknown random variables assumed to affect the variability of the data. The variances of the random-effects parameters, commonly known as variance components, become the covariance parameters for this particular structure. The MIXED procedure fits not only these traditional variance component models but numerous other covariance structures as well. In this case five standard covariance structures were fitted to the model; variance components, compound symmetry, first-order auto-regression, general, and toeplitz. These were compared using the 'By IC' method in line with the SAS user's guidelines produced by Kincaid (2005). This method compares the Akaike's Information Criteria (AIC) and log likelihood statistics of different models to identify the model which best fits the data.

This process was repeated for the internalising and externalising subscales. Results for the overall total difficulties score and externalising subscale indicated the use of the compound symmetry structure and results for the internalising subscale indicated the use of the first order auto-regression structure.

4.2.7.2 Testing the model

Though it was anticipated that the data may be positively skewed, as indicated by UK population norms (section 3.2.3.2), models were first run on the raw data to explore this. Diagnostic tests and plots were carried out on the selected models including a scatter plot of the residuals which allows for the visual assessment of the distance of each observation from the fitted line, a histogram with normal density which allows for visual assessment of the assumption that the measurement errors in the response variable are normally distributed, a Q-Q plot was used to compare the data set to the theoretical model, providing a graphical assessment of goodness of fit, and finally summary statistics for the residuals and the model fit (Figure 4.4).
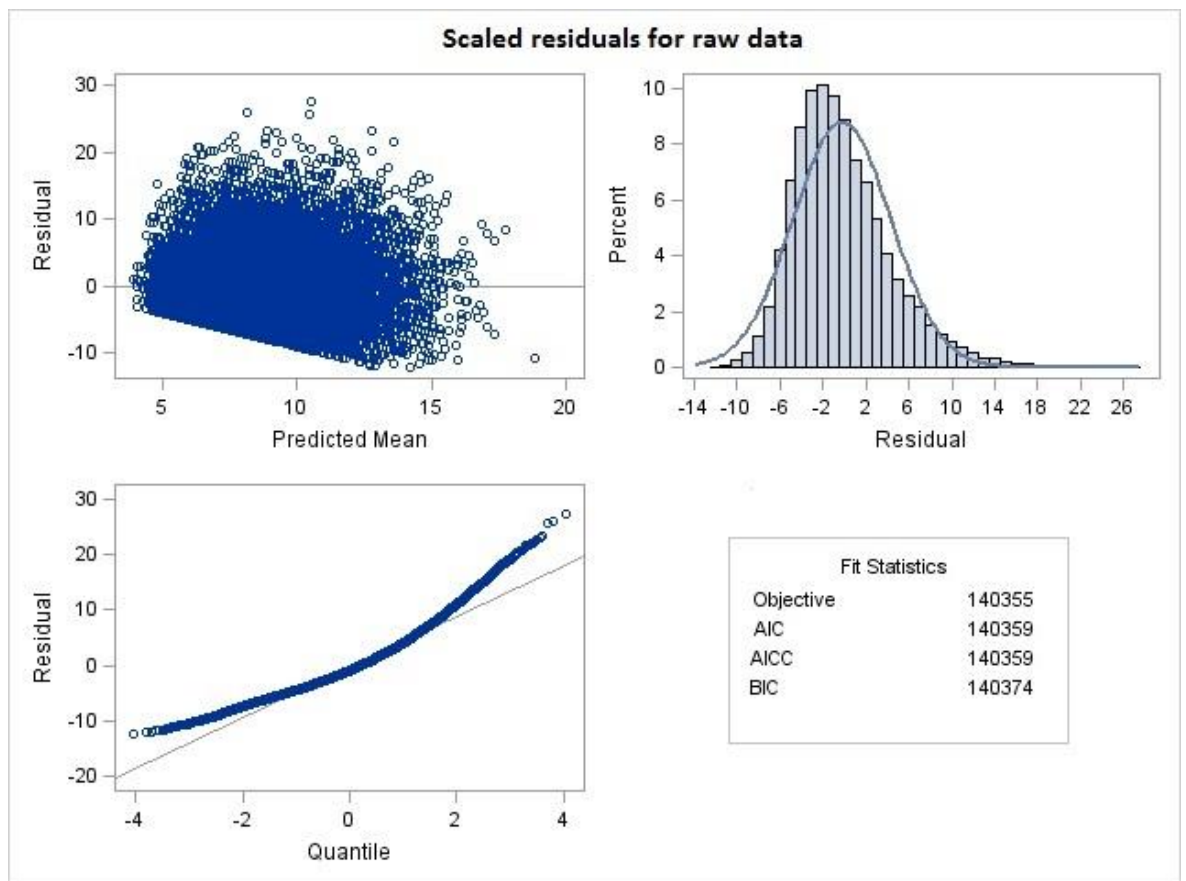
Figure 4.4. Multilevel modelling of the raw total difficulties data: Diagnostic plots.

Given the skew of the data, modelling the raw data resulted in high AIC values, indicating relatively poor fit of the data to the models. For statistical analysis it is optimal that the data be distributed normally with uniform variances between groups. Often when raw data is skewed transformations can be applied in order to stabilise the variance and normalise

skew. Methods of transforming the data were explored using the 'GLADDER' command in STATA, which revealed a logarithmic transformation as the most appropriate. In this instance natural log values to base *e* were generated for the outcome data. As log transformations will only work for positive values, before transformation could be carried out it was necessary to adjust the data further. According to Bland (2000, p320) the simplest approach to do this is to add a constant value to all observations before applying the logarithmic transformation. In this case adding a uniform value of '1' to all data entries was sufficient and made very little difference to the log transformed values. The model was then rerun using the log values as the outcome variable. The output and diagnostics were compared to those from the original model using the raw values. Although the fit to normal distribution was not perfect for the log values the transformed data appeared much closer to what would be considered a normal distribution than the raw data (Figure 4.5). The AIC value was also substantially lower for the model utilising the log values compared to that using the raw score indicating better fit of the data to this model.
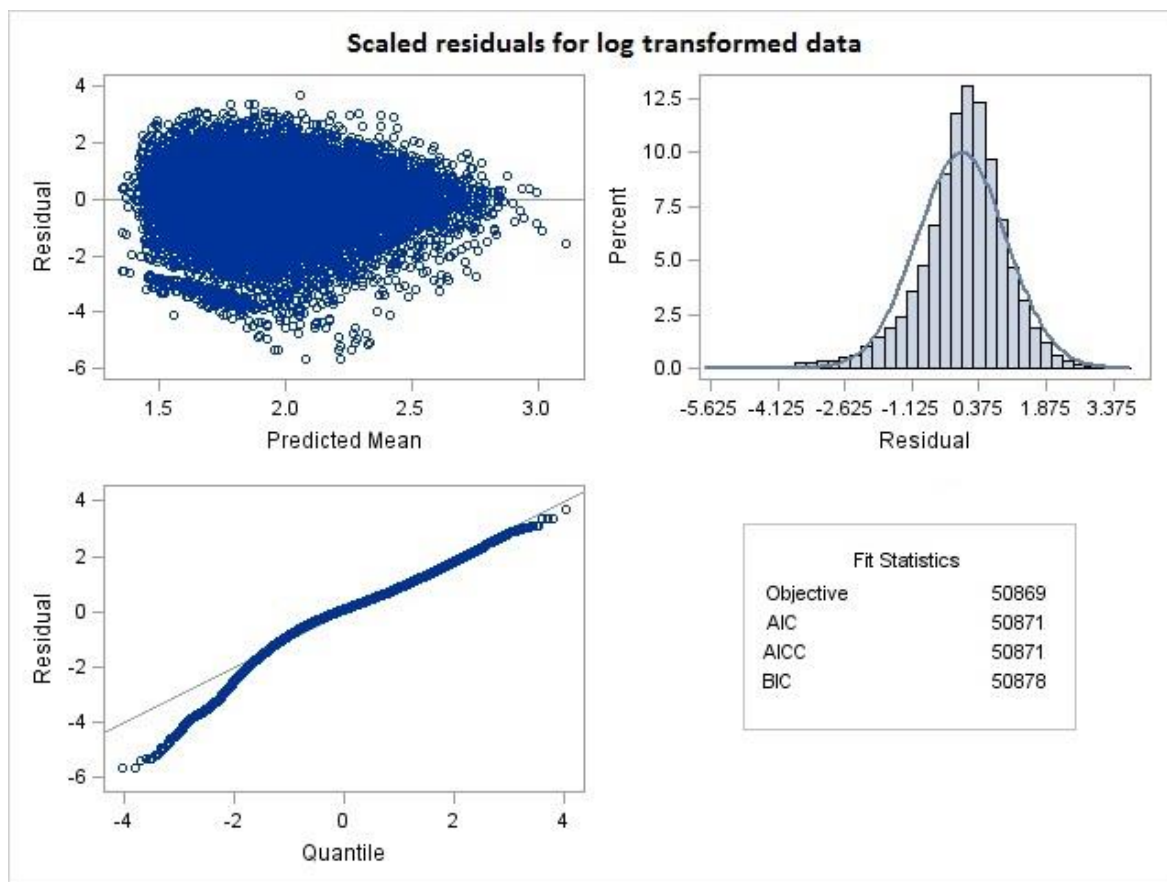
Figure 4.5 Multilevel modelling of the transformed total difficulties data: Diagnostic plots.

4.3 Results

4.3.1 Descriptive statistics

The demographic profiles for the women and their children at the primary data sweep are presented in Table 4.2 No major differences were identified between these demographic profiles, however, more women with undiagnosed depression compared to diagnosed depression reported being in poverty and being of an ethnic minority. As can be seen in Table 4.2, 17.8% of the undiagnosed group is comprised of non-white mothers compared to only 7% of the treated group and 9.9% of the control group. Similarly the proportions of families living within the OECD defined poverty range were greater within the undiagnosed group with 41.8% being classified as 'in poverty' compared to 37.8% of the treated group and only 25.5% of the control group being in poverty. This suggests that both being non-white and low income are related to depression status within this sample.

Table 4.2. Participant demographic characteristics according to maternal mental health group

	Diagnosed N= 957	Undiagnosed N= 1020	Non-depressed N= 8916
Maternal characteristics			
Age at CM birth (years), mean	28.7	28	29.2
Qualification Level			
Level 1	10.0	10.4	7.2
Level 2	34.5	28.5	29.1
Level 3	15.8	15.6	15.1
Level 4	21.7	23.3	32.8
Level 5	1.78	3.0	4.1
Overseas Only	2.3	3.4	1.8
None	13.9	15.8	9.9
Lone parent			
Yes	19	19.7	12.1
No	81	80.3	87.9
In poverty			
Yes	37.8	41.8	25.5
No	62.2	58.2	74.5
First live birth			
Yes	36.5	38.5	43
No	63.5	61.5	57
In employment			
Yes	40.7	44.1	53
No	59.3	55.9	47
Ethnicity			
White	93.0	82.2	90.0
Mixed	0.4	0.9	0.7
Indian	1.5	4.6	1.9
Pakistani/ Bangladeshi	3.0	7.7	3.7
Black/ Black British	1.6	3.3	2.5
Other	0.5	1.5	1.1
Child characteristics			
Child age (months), mean	9.2	9.2	9.2
Gender			
Male	48.8	52.8	50.4
Female	51.2	47.2	49.6
Ethnicity			
White	91.4	81.2	88.7
Mixed	2.3	2.8	2.5
Indian	1.6	4.3	1.9
Pakistani/ Bangladeshi	2.9	7.4	3.7
Black/ Black British	1.6	3.3	2.4
Other	0.2	1.1	0.8

Values are unweighted percentages of the sample unless stated otherwise.

4.3.1.1 British Ability Scales (BAS)

For assessments of the indicator of cognitive development, language development, children were dichotomised as scoring within the lowest 10% of the population or not. This 10% reflects the children who perform the most poorly on the language assessments. At sweeps 2 and 3, language was assessed using the Naming Vocabulary subscale, whereas the Word Reading subscale was used at sweep 4.

Some differences were observed in the proportions of children scoring within the lowest 10% of the population between groups. As can be seen in Figure 4.5, consistently lower proportions of children with non-depressed mothers scored within the lowest 10th percentile over time compared with children of depressed mothers. At ages 3 and 5 a greater proportion of the children with an undiagnosed depressed mother scored within the lowest 10th percentile compared to children of diagnosed depressed mothers (14.6% versus 9.2% at age 3, 12.6% versus 9.1% at age 5). At age 7 the proportions of children falling within the lowest 10th percentile were almost equal at 14.5% (diagnosed) and 14.4% (undiagnosed). At sweeps 2 and 3, which both used the Naming Vocabulary subscale, proportions of children scoring in the lowest 10% remain relatively stable. Slightly higher proportions of children in each of the three groups appear to score in the poorest 10% on the Word Reading measure at sweep 4. This is likely to reflect the fact that children are required to read words presented in text rather than name pictures presented to them which reflects a slightly more difficult task requiring slightly different cognitive functioning.
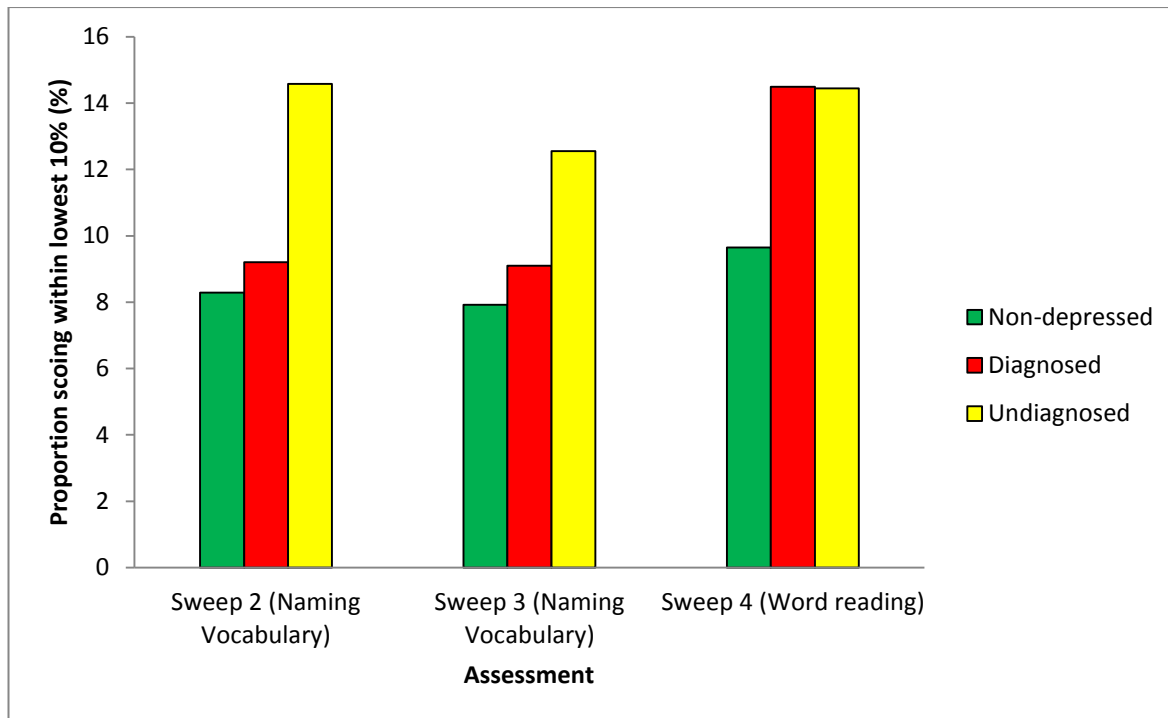
Figure 4.6. Proportions of children scoring in the lowest 10th percentile on the language scales of the BAS by maternal diagnosis group

The proportions of children scoring in the lowest 10% appeared to differ according to gender (Figure 4.7). Girls in all groups were generally less likely to scoring in the lowest 10%. Though there is little difference according to whether maternal depression was diagnosed or undiagnosed for boys, girls performance appears to be most associated with undiagnosed depression, however, these serve only as descriptive statistics and any associations do not account for potential confounders. These will be explored in the multilevel model.
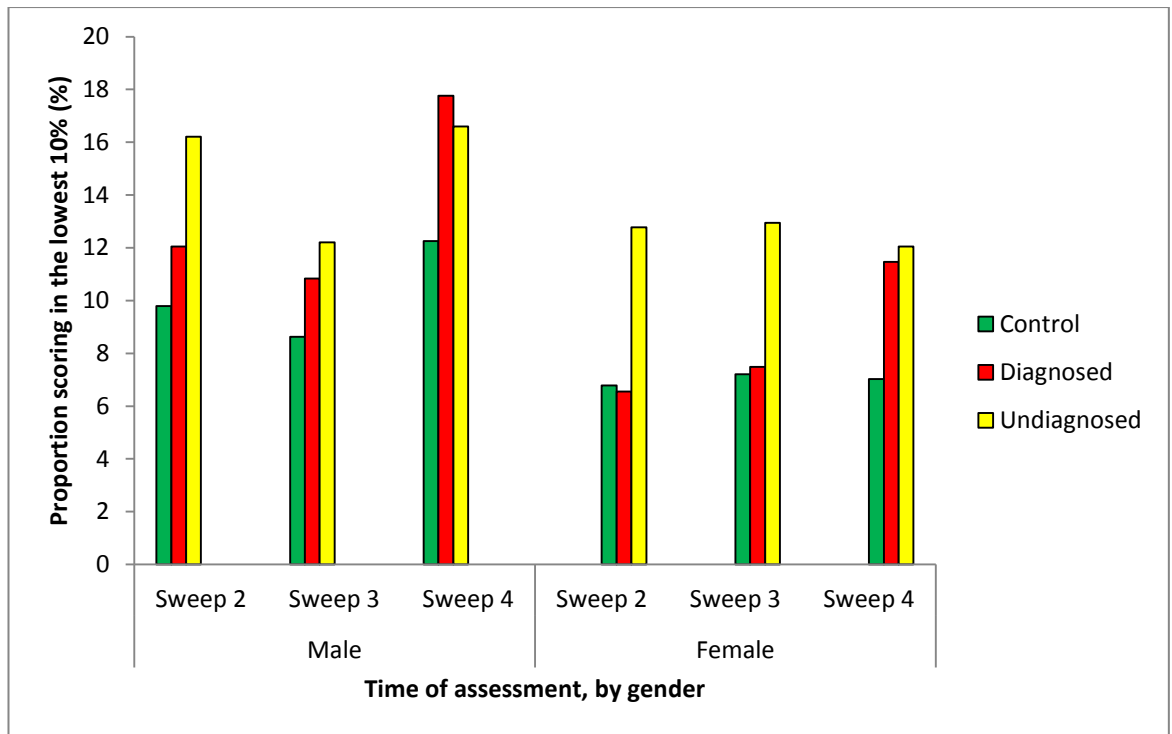
Figure 4.7 Proportions of children scoring in the lowest 10th percentile on the language scales of the BAS by maternal diagnosis group by gender

4.3.1.2 Strengths and Difficulties Questionnaire (SDQ)

Higher SDQ scores reflect greater levels of behavioural difficulties. Children were assessed at age 3 (sweep 2), age 5 (sweep 3) and age 7 (sweep 4). The mean SDQ scores between groups showed a tiered effect with children from the undiagnosed depressed group having the highest mean score at all three time points (11.5 points, 9.3 points and 9.4 points), children from the diagnosed group having slightly lower scores at all three time points (10.1 points, 8.3 points, and 9 points), and finally children from the control group having the lowest scores at each time point (8.4 points, 6.3 points, and 6.5 points) (see Figure 4.8).
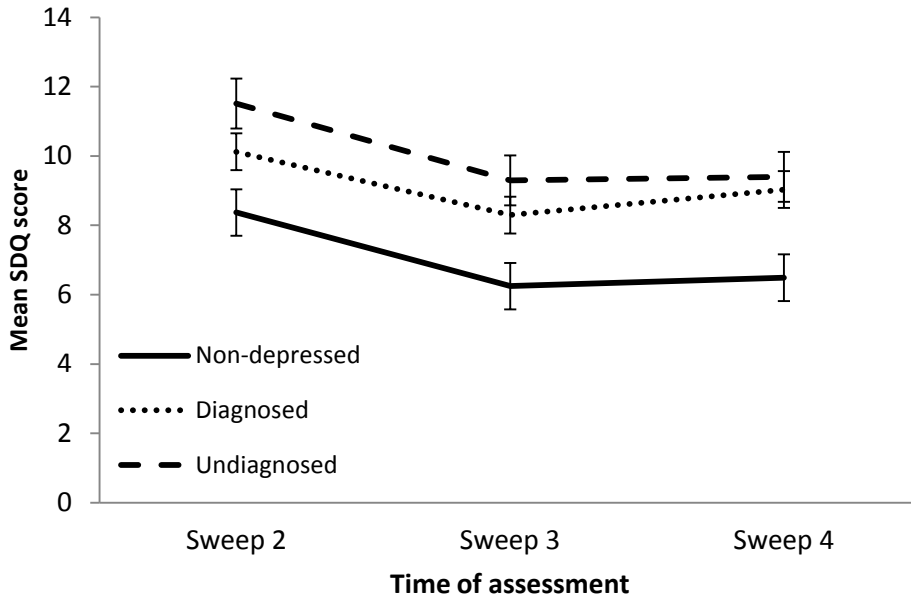
Figure 4.8. Mean child SDQ scores over time, by maternal mental health group at sweep 1

A similar pattern to that observed for the total difficulties score can be seen with the externalising subscale (Figure 4.9). Again a tiered effect of means between groups was found again with the children of untreated mothers scoring consistently higher (8.5, 6.5, 6.5), children of diagnosed mothers scoring just less (7.7, 5.7, 5.9), and children of control mothers scoring the lowest (5.9, 4.2, 4.2).
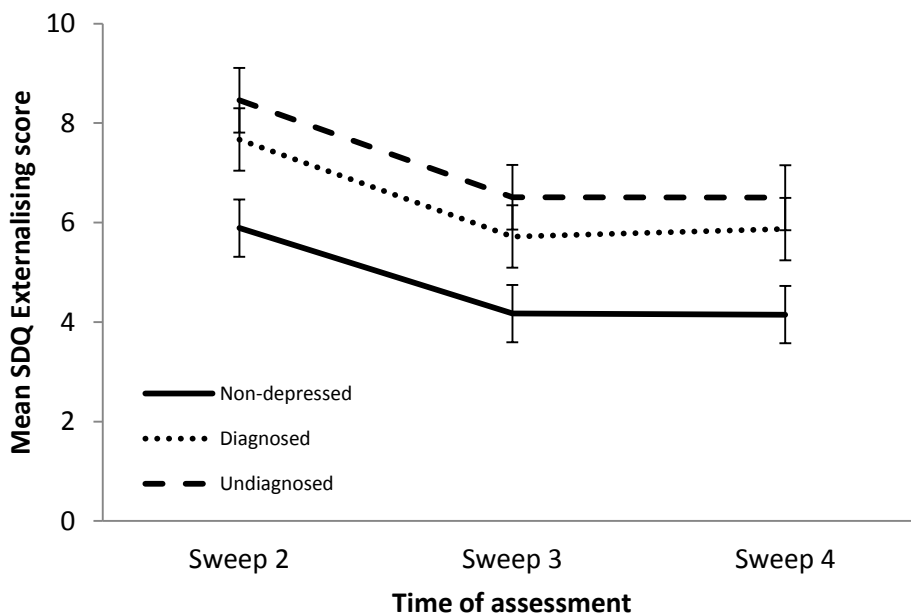
Figure 4.9. Mean SDQ externalising scores over time, by maternal mental health group at sweep 1

For the internalising subscale a different data trend is observed (Figure 4.10). Unlike the externalising and total scores a large decrease between sweep 2 and sweep 3 is not present. All group means remain relatively stable over time with the control group children showing the least variation (2.3, 2.0, 2.2). Both the children of diagnosed mothers and the untreated mothers exhibited the same pattern with a slight decrease in the mean score at sweep 3 compared to sweep 2 followed by an increase between sweeps 3 and 4.
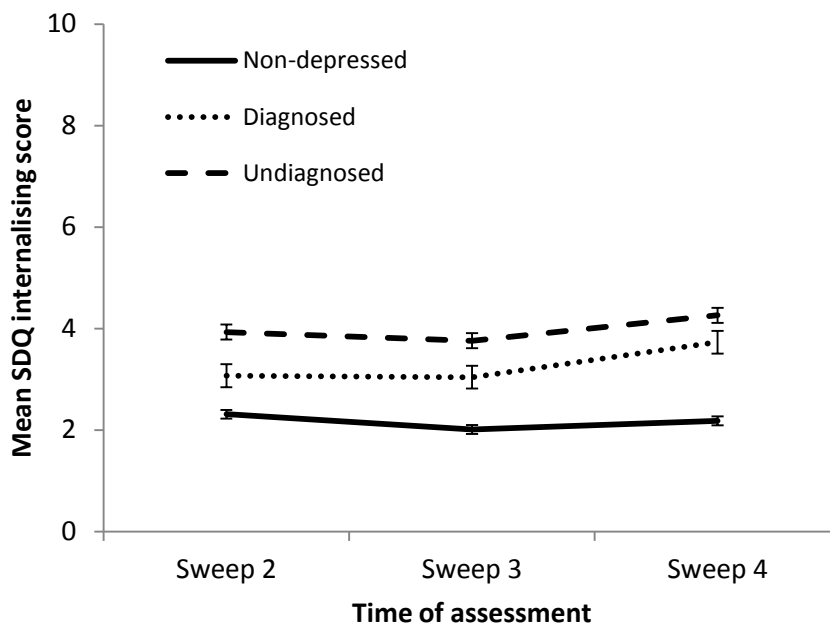
Figure 4.10. Mean SDQ internalising scores over time, by maternal mental health group at sweep 1

Mean total difficulties scores for the female cohort children were consistently lower at all time points by an average of 1.5 points, than those of male children irrespective of group (Figure 4.11). There appears to be some group related interaction as the differences between males and females vary between groups. The average difference between control males and females over time was .95 points, compared to 1.46 points for the diagnosed group, and 1.95 points for the undiagnosed group.
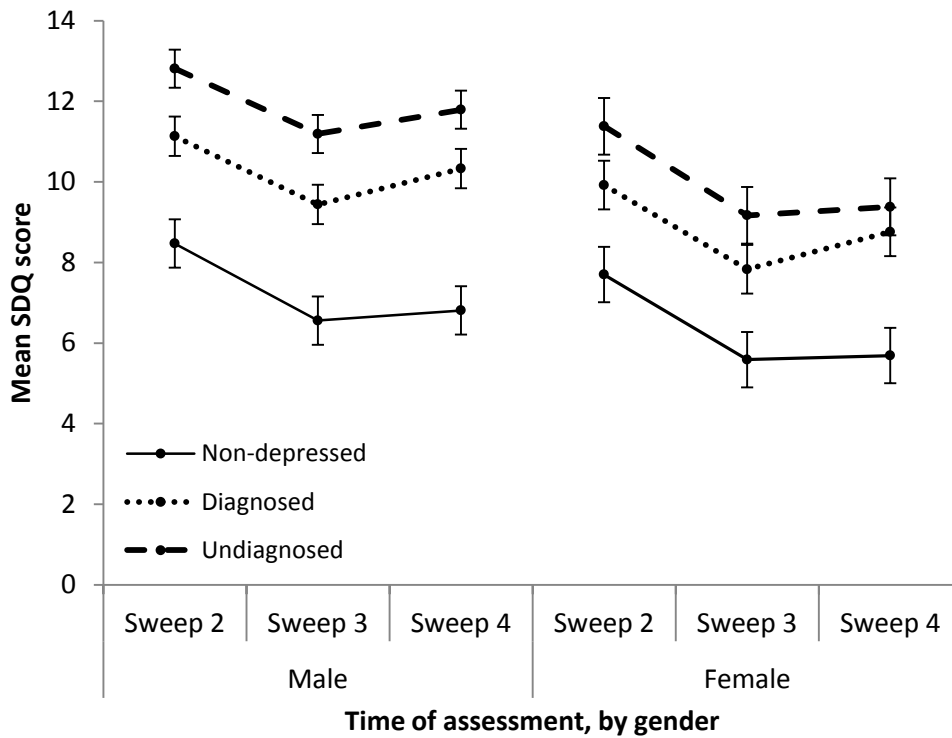
Figure 4.11. Mean SDQ scores for male and female children over time, by maternal mental health group at sweep 1

As with the total difficulties score, mean externalising scores for the female cohort children were consistently lower at all time points, than those of male children irrespective of group. Again a similar group and gender interaction is observed with the average difference between control males and females over time being .93 points, compared to 1.30 points for the diagnosed group, and 1.41 points for the undiagnosed group.

Figure 4.12. Mean SDQ externalising scores for male and female children over time, by maternal mental health group at sweep 1

Mean scores for all groups were fairly low. The greatest mean group score was observed in male children of undiagnosed mothers (4.54) at sweep 4, with the lowest mean group score owing to female children of control mothers (1.99) at sweep 3. Only very small differences were observed between genders for the internalising subscale with the average difference between control males and females over time being .08 points, compared to .32 points for the diagnosed group, and .37 points for the undiagnosed group. Despite these differences being small they do follow the same pattern as the results for the externalising subscale and the total difficulties scale.
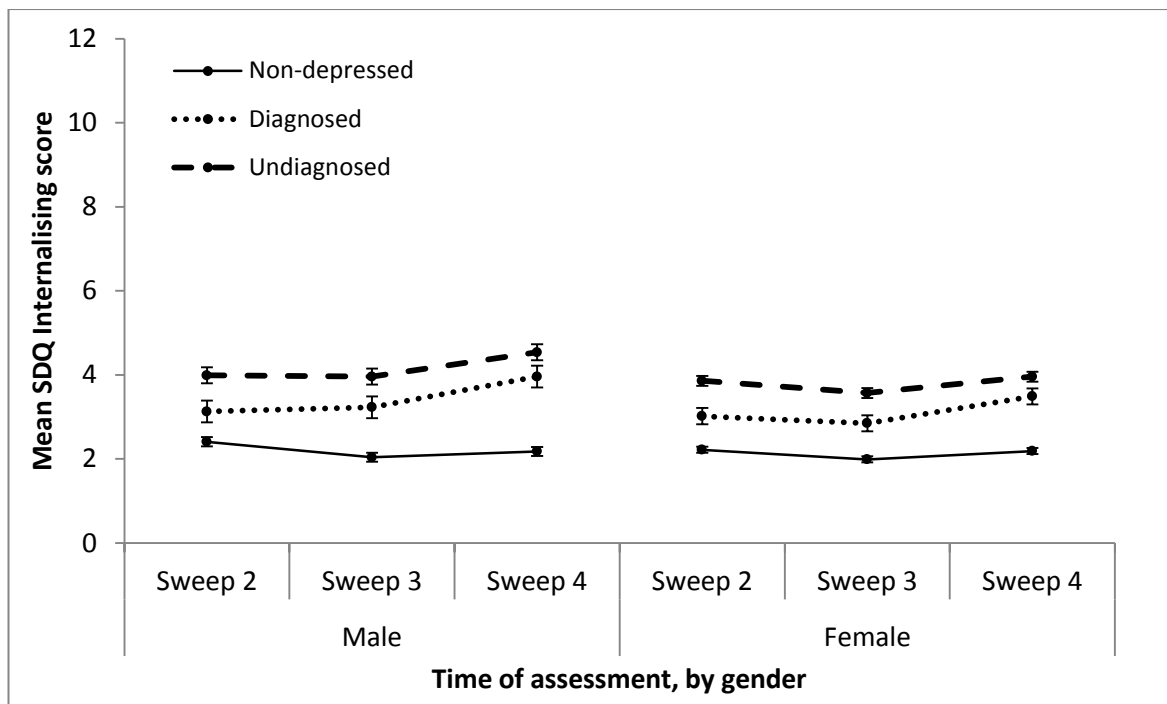
Figure 4.13. Mean SDQ internalising scores for male and female children over time, by maternal mental health group at sweep 1

For all groups a change in mean scores was observed over time, with improvements being observed as children get older. This is likely due to natural behaviour improvements associated with age.

4.3.1.3 Missing data

Proportions of missing data were examined for each of the group according to each of the covariates and each of the outcome measures. As can be seen in Table 4.3, proportions of missing data for each variable were relatively similar between groups. On the whole levels of missing data for the covariates were low (<1%) with the exception of poverty data whereby 7% of data was missing overall.

For behavioural outcome data a slight increase in missing data was noted for both groups of depressed mothers compared to non-depressed mothers at each of the three sweeps. This is particularly pronounced when comparing the numbers of children for whom no SDQ data is available with the rate of completely missing outcome data being almost double in the undiagnosed group compared to the diagnosed group (Table 4.4).

Much smaller differences were observed between groups for the language development, with all groups having around 5% of outcome data completely missing.

Table 4.3. Missing covariate data at sweep 1 by group

	All N (%)	Non-depressed N (%)	Diagnosed N (%)	Undiagnosed N (%)
Maternal characteristics				
Age at birth	4 (0.04)	4 (0.04)	0 (0)	0 (0)
Qualifications	7 (0.06)	5 (0.06)	1 (0.10)	1 (0.10)
Lone parent status	0 (0)	0 (0)	0 (0)	0 (0)
First live birth	13 (0.12)	3 (0.03)	8 (0.84)	2 (0.20)
Employment status	5 (0.04)	3 (0.03)	0 (0)	2 (0.20)
Ethnicity	21 (0.19)	20 (0.22)	1 (0.10)	0 (0)
Poverty status	823 (7.37)	629 (7.05)	52 (5.43)	76 (7.45)
Child characteristics				
Age at Sweep1	0 (0)	0 (0)	0 (0)	0 (0)
Gender	0 (0)	0 (0)	0 (0)	0 (0)
Ethnicity	21 (0.19)	20 (0.22)	1 (0.10)	0 (0)

Table 4.4. Missing outcome data by group

	All N (%)	Non-depressed N (%)	Diagnosed N (%)	Undiagnosed N (%)
Behavioural development (SDQ)				
Sweep 2	3501 (32.14)	2737 (33.39)	337 (35.21)	427 (41.86)
Sweep 3	2283 (20.96)	1754 (21.40)	236 (24.66)	293 (28.73)
Sweep 4	1927 (17.69)	1471 (17.95)	190 (19.85)	266 (26.08)
All missing	660 (5.90)	448 (5.02)	62 (6.48)	98 (9.61)
Language development (BAS)				
Sweep 2	527 (4.84)	423 (5.16)	44 (4.60)	60 (5.80)
Sweep 3	599 (5.50)	471 (5.75)	56 (5.85)	72 (7.06)
Sweep 4	793 (7.28)	634 (7.74)	67 (7.00)	92 (9.02)
All missing	527 (4.84)	423 (5.16)	44 (4.60)	60 (5.80)

As the multiple imputation process uses children's scores at alternative time points to predict missing time points, scores could not be imputed for those children who had missing time data at all three time points. Consequently the final imputed data sets provide complete behavioural outcome data for 10, 502 children (8468 non-depressed mothers, 895 diagnosed mothers, 922 undiagnosed mothers), and complete language outcome data for 10, 635 children (8493 non-depressed mothers, 913 diagnosed mothers, 960 undiagnosed mothers).

4.3.2 Multilevel modelling

4.3.2.1 Formulating the model

The correlations between the variables to be used in the multilevel models were assessed prior to analysis. These correlations are presented in Table 4.5. As can be seen, a weak to moderate correlation was found between income and employment status and income and being a lone parent. These are shown to be insignificant. A very strong correlation was identified between the mother's ethnicity and the child's ethnicity as would be expected. Upon inspection of these variables the number of children whose ethnicity differed from the mother's was found to be very low (2.9%), the majority of which were classified as mixed race (69.2%). As ethnicity is an explanatory variable at both the mother and child level, both were included in the final analysis.

Table 4.5. Correlations between all variables included in the statistical models

Variables	SDQ score	Mother age	Mother Ethnicity	Child Ethnicity	Qualifications	Child Age	Parity	Employment	Lone Parent	Sex	Poverty	Group
SDQ score	1											
Mother age	-0.20*	1										
Mother Ethnicity	0.06*	-0.02*	1									
Child Ethnicity	0.07*	-0.03*	0.92*	1								
Qualifications	0.16*	-0.08*	0.22*	0.23*	1							
Child Age	-0.14*	0.00	0.00	0.00	0.00	1						
Parity	-0.02	0.30*	0.05*	0.06*	0.13*	-0.00	1					
Employment	0.17*	-0.19*	0.15*	0.16*	0.25*	-0.11*	0.09*	1				
Lone Parent	0.15*	-0.22*	0.03*	0.05*	0.13*	0.06*	-0.06*	0.15	1			
Sex	-0.11*	0.04	0.00	0.00	0.01	0.00	0.02*	0.00	0.00	1		
Poverty	0.22*	-0.29*	0.18*	0.20*	0.32*	-0.01*	0.10*	0.43*	0.42*	0.01*	1	
Group	0.27*	-0.08*	0.07*	0.07*	0.10*	0.00	0.04*	0.12*	0.12*	-0.01	0.16*	1

* $p<.05$

4.3.2.2 Fitting the model to the data

i. British Ability Scales: Vocabulary subscales

A multilevel model with a compound symmetry covariance structure was found to best represent the data using the method described above. This was used to assess whether children of mothers with differing mental health circumstances in the postnatal year were more or less likely to score within the lowest 10th percentile on the vocabulary scales of the British Ability Scales. Children's scores were dichotomised into those scoring in the lowest 10th percentile and those scoring above. Table 4.6 shows the results from the multilevel model.

Table 4.6. Multilevel analysis showing the association between maternal depression status and children's vocabulary scores over time.

Comparison	OR	95% CI	Pr > |t|
Control and Undiagnosed			
Sweep 2	1.19	0.88-1.75	0.1544
Sweep 3	1.16	0.85-1.71	0.2014
Sweep 4	1.00	0.61-1.48	0.9916
Control and Diagnosed			
Sweep 2	0.85	0.52-1.26	0.2573
Sweep 3	0.99	0.54-1.56	0.9657
Sweep 4	1.15	0.75-1.63	0.2135
Diagnosed and Undiagnosed			
Sweep 2	1.40	1.12-1.85	0.0584
Sweep 3	1.16	0.81-1.56	0.3429
Sweep 4	0.87	0.60-1.15	0.3403

As can be seen from Table 4.6 no significant differences were found in the model investigating the relationship between maternal mental health diagnosis and child cognitive problems. All odds ratios were very close to 1 and were all insignificant. There appeared to be a slight difference between children of diagnosed and undiagnosed mothers at sweep 2 which was almost significant ($p = .06$) but given the size of the difference and the other findings this is likely to be unimportant. As a significant association between maternal depression and child cognitive development was not expected based on previous research using this sample, multiple imputation was not deemed necessary for this model.

ii. SDQ : Externalising behaviours construct

Using the method above, a multilevel model with a compound symmetry covariance structure was used to explore the differential effects of undiagnosed and diagnosed maternal depression on child externalising behaviour development, a composite construct of only two subscales (hyperactivity/ inattention and conduct problems). Table 4.5 shows the results from the multilevel model carried out using the data set.

Small but significant differences were found between the children of undiagnosed mothers and children of control mothers with children of undiagnosed mothers scoring 0.24 points more than children of control mothers at age three, 0.32 points more at age five, and 0.27 points more at age seven. These differences were all were highly statistically significant (p <.0001). Similarly differences were found between the children of diagnosed mothers and children of control mothers with children of undiagnosed mothers scoring 0.12 points more than children of control mothers at age three, 0.18 points more at age five, and 0.21 points more at age seven. These differences were also all highly statistically significant (p <.0001).

Small differences were again found between the children of diagnosed and undiagnosed mothers, with a significant difference being found at ages 3 and 5 but not at age 7.

No effect of missing data was observed as shown by the very slight differences between actual and imputed values presented in Table 4.7 No changes to significance were made thus missing data is unlikely to have an effect.

Table 4.7. Multilevel analysis showing the association between maternal depression status and children's externalising behaviours scores over time

Comparison	Coefficient	Lower	Upper	Pr > |t|
Control and Undiagnosed				
Sweep 2	0.24	0.18	0.30	<.0001
	0.21	*0.16*	*0.26*	*<.0001*
Sweep 3	0.32	0.26	0.38	<.0001
	0.29	*0.24*	*0.35*	*<.0001*
Sweep 4	0.27	0.21	0.32	<.0001
	0.25	*0.19*	*0.31*	*<.0001*
Control and Diagnosed				
Sweep 2	0.12	0.08	0.20	<.0001
	0.12	0.07	0.17	*<.0001*
Sweep 3	0.18	0.13	0.24	<.0001
	0.16	0.11	0.22	*<.0001*
Sweep 4	0.21	0.16	0.27	<.0001
	0.20	0.14	0.25	*<.0001*
Diagnosed and Undiagnosed				
Sweep 2	0.09	0.02	0.16	0.0102
	0.08	0.02	0.14	*0.0141*
Sweep 3	0.11	0.05	0.18	0.0006
	0.11	0.05	0.18	*0.0005*
Sweep 4	0.05	-0.01	0.11	0.1254
	0.04	-0.02	0.12	*0.1605*

Model accommodated for sampling weights.
Imputed values

iii. SDQ: Internalising behaviours construct

For the internalising subscale a multilevel model with the first order autoregressive covariance structure was used. The results from this multilevel model are demonstrated in Table 4.8. As with the externalising scores, small but significant differences were found between the children of undiagnosed mothers and children of control mothers with children of undiagnosed mothers scoring 0.32 points more than children of control mothers at age three, 0.34 points more at age five, and 0.31 points more at age seven. These differences were all highly statistically significant (p <.0001). Similarly minor differences were found between the children of diagnosed mothers and children of control mothers with children of undiagnosed mothers scoring 0.12 points more than children of control mothers at age three, 0.19 points more at age five, and 0.25 points more at age seven. These differences were also all highly statistically significant (p <.001). Significant differences were found between the children of diagnosed and undiagnosed mothers at age three with children of

undiagnosed scoring 0.19 points more than children of diagnosed mothers, and at age five, where children of undiagnosed mothers scores 0.12 points more than children of diagnosed mothers (p<.001). At age seven this difference was no longer significant (p =.1840).

Again, no obvious effect of missing data is observed.

Table 4.8. Multilevel analysis showing the association between maternal depression status and children's internalising behaviours scores over time

Comparison	Coefficient	Lower	Upper	Pr > |t|
Control and Undiagnosed				
Sweep 2	0.32	0.25	0.40	<.0001
	0.26	*0.19*	*0.32*	*<.0001*
Sweep 3	0.34	0.28	0.41	<.0001
	0.30	*0.24*	*0.37*	*<.0001*
Sweep 4	0.31	0.25	0.37	<.0001
	0.28	*0.22*	*0.35*	*<.0001*
Control and Diagnosed				
Sweep 2	0.12	0.05	0.17	0.0002
	0.10	*0.04*	*0.15*	*0.0003*
Sweep 3	0.19	0.14	0.25	<.0001
	0.18	*0.12*	*0.24*	*<.0001*
Sweep 4	0.25	0.19	0.32	<.0001
	0.22	*0.17*	*0.28*	*<.0001*
Diagnosed and Undiagnosed				
Sweep 2	0.19	0.12	0.29	<.0001
	0.15	*0.07*	*0.23*	*<.0001*
Sweep 3	0.12	0.05	0.20	0.0005
	0.10	*0.03*	*0.18*	*0.0055*
Sweep 4	0.05	-0.02	0.12	0.1840
	0.05	*-0.02*	*0.12*	*0.1483*

Model accommodated for sampling weights.
Imputed values

iv. SDQ: Total difficulties score continuous

A compound symmetry covariance structure was indicated for the model exploring the total score. Table 4.9 shows the coefficients, 95% confidence intervals and p-values derived from the multilevel model.

As would be expected given the similar patterns of results shown in the SDQ subscales, the results for the total behaviour score further reflected these findings. Children of depressed mothers (undiagnosed or diagnosed) performed consistently significantly worse than children of non-depressed mothers. Net of other predictive variables, children from the undiagnosed group score 0.26 points (95% CI, 0.19 to 0.33) more than children of non-depressed mothers at age three, p <.0001, 0.37 (CI, 0.31 to 0.43 points) points more at age 5, p <.0001, and 0.31 points (CI, 0.25 to 0.37 points) more at age 7, p <.0001. Point differences were slightly smaller for children of diagnosed mothers, 0.14 points (95% CI, 0.08 to 0.20 points) at age 3, p =.0003, 0.24 (CI, 0.18 to 0.30 points) points more at age 5, p <.0001 and 0.27 points (CI, 0.22 to 0.33 points) more at age 7, p <.0001.

When comparing the children of undiagnosed and diagnosed mothers, children of undiagnosed mothers were shown to perform consistently worse than children of diagnosed mothers, though this was only significant at ages 3 and 5. The difference in effect sizes decreased over time between these groups.

Table 4.9. Multilevel analysis showing the association between maternal depression status and children's SDQ total difficulties scores over time

Comparison	Coefficient	Lower	Upper	Pr > |t|
Control and Undiagnosed				
Sweep 2	0.26	0.19	0.33	<.0001
	0.24	*0.18*	*0.29*	*<.0001*
Sweep 3	0.37	0.31	0.43	<.0001
	0.33	*0.27*	*0.39*	*<.0001*
Sweep 4	0.31	0.25	0.37	<.0001
	0.30	*0.23*	*0.36*	*<.0001*
Control and Diagnosed				
Sweep 2	0.14	0.08	0.20	<.0001
	0.13	*0.08*	*0.18*	*<.0001*
Sweep 3	0.24	0.18	0.30	<.0001
	0.20	*0.15*	*0.26*	*<.0001*
Sweep 4	0.27	0.22	0.33	<.0001
	0.24	*0.18*	*0.30*	*<.0001*
Diagnosed and Undiagnosed				
Sweep 2	0.12	0.03	0.19	0.0055
	0.10	*0.03*	*0.16*	*0.0023*
Sweep 3	0.10	0.04	0.18	0.0023
	0.12	*0.04*	*0.17*	*0.0007*
Sweep 4	0.03	-0.03	0.10	0.3376
	0.05	*-0.02*	*0.12*	*0.1870*

Model accommodated for sampling weights.
Imputed values

v. SDQ: Total difficulties score dichotomised

The final multilevel model adopting the compound symmetry structure explored the likelihood of children scoring above threshold for behavioural difficulties according to their mothers' mental health diagnosis. Thresholds were derived by identifying the poorest performing 10% of the sample in line with the SDQ guidance discussed in section 3.2.3.2. This yielded a cut-off point of 14 points and above. Table 4.10 shows the odds ratios, 95% confidence intervals and p-values derived from the multilevel model.

Children of undiagnosed mothers were shown to be more than twice as likely as children of non-depressed mothers to score above threshold for behavioural difficulties at all three time points. At age three children of undiagnosed depressed mothers were 2.4 times as likely to be classified as having behavioural difficulties compared to children of non-depressed mothers, at age five they were 3.2 times as likely and at age seven they were 2.3 times as likely. These odds ratios were all shown to be highly significant (p<.0001).

Similar results were observed between the diagnosed and control groups with children of diagnosed depressed mothers again been around twice as likely as controls to score above threshold for behavioural difficulties. At age three children of diagnosed depressed mothers were 1.8 times as likely to be classified as having behavioural difficulties compared to children of non-depressed mothers, at age five they were 2.3 times as likely and at age seven they were 2.3 times as likely. These odds ratios were again all shown to be highly significant (p<.0001).

Small differences were observed in the comparison between children of diagnosed versus children of undiagnosed mothers. At age three, children of undiagnosed mothers were 1.34 times as likely to score above threshold for behavioural difficulties, this rose slightly at age five to 1.37 times as likely. These differences both just reached statistical significance (p<.05). At age seven there was no longer an increased likelihood of a child from an undiagnosed mother scoring above threshold for behavioural difficulties compared to a child from a diagnosed mother. Children of both types of mother were equally likely to score above threshold for behavioural difficulties by age seven.

Table 4.10. Multilevel analysis showing the association between maternal depression status and children's SDQ total difficulties scores over time

Comparison	OR of difficulties	Lower	Upper	Pr > |t|
Control and Undiagnosed				
Sweep 2	2.41	1.97	2.95	<.0001
Sweep 3	3.17	2.59	3.87	<.0001
Sweep 4	2.31	1.91	2.78	<.0001
Control and Diagnosed				
Sweep 2	1.80	1.47	2.21	<.0001
Sweep 3	2.30	1.85	2.87	<.0001
Sweep 4	2.30	1.90	2.78	<.0001
Diagnosed and Undiagnosed				
Sweep 2	1.34	1.02	1.75	0.0342
Sweep 3	1.37	1.05	1.80	0.0204
Sweep 4	1.00	0.79	1.28	0.9853

Model accommodated for sampling weights.

4.3.2.3 Interactions

i. Child gender

Interaction terms between the grouping and child gender variables were added into the multilevel models to explore whether boys and girls are differently affected by their mother's depression status in terms of their cognitive and behavioural outcomes. No evidence of a significant interaction between gender and maternal depression group was identified for the cognitive development model (p=0.3121). Similarly no evidence of a significant interaction was found for the behavioural model measured using total difficulties scores (p=.4171), the externalising subscale model (p=.5299) or on the internalising subscale model (p=.1048).

ii. Child Ethnicity

Interaction terms between the grouping and ethnicity variables were added into the multilevel models to explore whether children of different ethnicities are differently affected by their mother's depression status in terms of their cognitive and behavioural outcomes. Again, no evidence of a significant interaction between ethnicity and maternal depression group was found for the cognitive development model (p=0.1504) or the behavioural model measured using total difficulties scores (p=.5961). Similarly there was no evidence of a significant interaction having effect on the externalising subscale (p=.7037) or the internalising subscale (p=.6146).

4.3.3 Testing of model assumptions

Diagnostic tests and plots were carried out for each model exploring behavioural development to test whether the data meets the model assumptions (see appendix 3). These plots allow us to discern whether there are non-linearities or unequal error variances in the data. A non-constant spread of the residuals, such as a tendency for more clustered residuals for small fixed values and more dispersed residuals for large fixed values would indicate heteroscedasticity. In this instance no particularly apparent trends or patterns are obvious indicating that the data is not subject to these issues. Histograms of the residuals revealed slight negative skews as the mass of the distribution is concentrated to the right of the figure. This means that there are relatively few low values. In all, the diagnostic tests suggest that the data meets the model assumptions relatively well and that the data is a good fit to the model.

4.4 Discussion

The analyses in this chapter were undertaken to explore the effect of failing to diagnose maternal depression during the first postnatal year on child cognitive and behavioural development over time. Although the adverse effects of diagnosed maternal depression on child development have been well documented, less is known about the effect undiagnosed maternal depression has on children. This analysis is the first which has compared the longitudinal outcomes in children of mothers with self-reported symptoms of depression

but who were not clinically diagnosed, with both children of control mothers and children of mothers who had been diagnosed and were receiving treatment for depression. Multilevel modelling was used to draw comparative estimates between groups for both measures of cognitive and behavioural development. The verbal component of the British Ability Scales (Elliot, 1983) was used as a measure of cognitive development with the SDQ (Goodman, 1997) being used as the measure of behavioural development.

4.4.1 Principal findings

The present analysis identified 8.6% (n= 957) of the sample population as having been diagnosed and receiving treatment for depression and 9.1% (n= 1020) as having symptoms of depression according to a self-rated scale but not receiving any treatment for it at the nine month survey. The primary objective of the analysis was to explore the effect of undiagnosed postnatal depression on child cognitive and behavioural development over time.

4.4.2 Child cognitive development; language development

The results of the multilevel analysis showed no indication of depression being related to poor performance on this measure, either diagnosed or undiagnosed, at any time point. It therefore appears that postnatal depression does not significantly affect children's language development, based on word naming and reading assessments, in this sample.

4.4.3 Child behavioural development

The results of the analysis investigating behavioural development, assessed using the SDQ, showed small but statistically significant associations between maternal depression, both diagnosed and undiagnosed, and children's behavioural scores up to the age of seven years with children of depressed mothers performing consistently more poorly than children of non-depressed mothers. In addition to scoring higher on the SDQ scale, reflective of higher levels of behaviour problems, children of depressed mothers were more than twice as likely as children of non-depressed mothers to score above threshold for behavioural problems.

Differences were also observed between the children of undiagnosed and diagnosed mothers. Children of mothers with undiagnosed depression performed significantly worse than children of diagnosed mothers at ages three and five but not at age seven. These differences were observed both in terms of point scoring and the likelihood of scoring above threshold.

4.4.4 Other findings

One of the secondary aims of the analysis was to investigate the impact of undiagnosed depression on individual facets of behaviour, specifically, externalising and internalising behaviour problems. The actual differences in effect sizes were again very small between groups, though those for the externalising scores were slightly greater suggesting maternal depression may be more closely related to conduct disorders and inattention/hyperactivity issues than emotional and peer related problems.

The final aim of the analysis was to identify whether child gender or ethnicity mediates the effect of maternal depression according to whether or not the mother was diagnosed. No significant effects related to child gender or ethnicities were identified when an interaction term was applied to maternal group allocation.

4.4.5 Comparison with other literature

Literature exploring the effects of maternal depression on child cognitive development shows conflicting findings with some authors reporting that particular aspects of cognitive development are affected by maternal depression (Murray, 1992), whilst others report no differences (Kiernan and Mensah, 2009). In the most recent systematic review of the literature, Grace et al. (2003) draw on the findings of 7 studies reporting on the relationship between maternal depression and child cognition as discussed in Chapter 1. Although minor effects on certain aspects of cognition such as object recognition were found, no significant differences were found relating to broader measures of general cognition. Conversely, in more recent research, a significant association between postnatal depression and language development has been identified (Stein et al., 2008). This example highlights the disparity in findings between studies in this field and the need for more conclusive evidence.

The results from the present analysis show no significant association between postnatal depression and child language development. This was carried out on a far greater, nationally representative sample of 11,162 mother-child pairs, compared to the much smaller sample of 999 mother-child pairs included in the study conducted by Stein and colleagues. The present analysis also examines children over a longer period of time, to age 7, thus allowing for any temporal changes in effects to be observed. In this case no significant differences were found at any time point adding to the body of evidence that does not provide support for an association between postnatal depression and child cognitive development.

Whilst the literature surrounding the effects of postnatal depression on child cognitive development is varied in its outcome, studies based on exploring the effects on child behavioural development have been more consistent, generally agreeing that there is a clear association.

Although the actual effect sizes found in the present analysis were very small, they were statistically significant which would usually indicate that the differences are unlikely to be due to chance. However, there has been much discussion regarding statistical versus clinical significance, as some authors have argued that even though a difference may be statistically significant, it may not actually represent a meaningful clinical difference (Sterne and Smith, 2001, Houle and Stump, 2008). Previous work has indicated that a 1 point difference in SDQ scores is enough to bring about an increased likelihood of disorder (Goodman and Goodman 2009). In this study it was found that the odds of a child being diagnosed with a behavioural disorder were shown to increase at a constant rate across the range of scores (OR 1.28 per one-point increase in the SDQ scores). In the present analysis statistically significant effect sizes were shown to be less than 1 point, ranging between 0.1 points and 0.37 points. No previous literature has indicated whether such small differences were likely to be clinically meaningful. As a consequence of this, the continuous scores were dichotomised into 'low' and 'high' risk groups for further analysis using the method described in Goodman (2001). Similar patterns were observed using the dichotomous data as the outcome as were found for the continuous data with children of diagnosed or undiagnosed mothers being more than twice as likely as controls to fall into the 'high' risk

category, and children of mothers with undiagnosed depression being more likely to be 'high' risk than children of diagnosed mothers until age five.

The importance of mothers' mental health in relation to child behavioural development has been highlighted in small and large scale studies, reviewed by Goodman et al., (2011) and Grace et al., (2003). Findings from the present analysis are in line with this as children of both diagnosed and undiagnosed depressed mothers scored consistently worse than controls. The results of the present analysis go further to build on the findings of previous work by additionally exploring how failing to be diagnosed and treated for depression affects the impact on child behaviour.

The MCS has previously been used to investigate the interaction between maternal mental health and child development. For instance, Kiernan and Mensah (2009) explored the effect of maternal depression on child cognitive and behavioural development. After adjustment for confounding factors such as family income and child's age and gender, children of mothers scoring high on the Malaise Inventory (4 or more points) are 1.71 times more likely to have a high score (14-40) on the strengths and difficulties questionnaire at age 3 years ($p<.001$). The present analysis builds on this methodology by not only looking at mothers who scored high for depressive symptoms on the Malaise Inventory but more specifically stratifying mothers into those who were in treatment for depression and those who scored high but were not in treatment. Despite this slightly different methodology the results of the present analysis were similar to those noted by Kiernan and Mensah. In the present analysis, children of mothers with undiagnosed depressive symptoms at 9 months were 2.4 times more likely than controls to have a high SDQ score (14-40) on the SDQ at age 3. In comparison children of mothers who were diagnosed and receiving treatment at 9 months were 1.8 times more likely than controls to have a high SDQ score at age 3. This shows children of all depressed mothers to be around twice as likely to score high on the SDQ in comparison to controls, a difference similar to that found by Kiernan and Mensah. Unlike the present analysis Kiernan and Mensah do not track changes over time but only at age 3, meaning that results at ages 5 and 7 years cannot be compared.

Previous studies have indicated that boys may react more severely than girls to maternal depression with boys of depressed mothers often scoring worse than girls of depressed

mothers on perceptual and motor tasks (Hay and Kumar, 1995) as well as behavioural assessments (Kiernan and Mensah, 2009, Sinclair and Murray, 1998). In the present analysis gender interaction terms were added to the models to investigate whether child gender was important with regards to diagnosis of depression. The results showed no effect related to gender.

4.4.6 Interpretation of findings

As discussed in Chapter 2, differences in attachment style and communication between mother and infant are the likely mechanism through which maternal depression affects children. Depressed mothers have been found to interact differently with children compared to non-depressed mothers, being less responsive and affectionate, less communicative with infants and sometimes even hostile and over-stimulating during play (Milgrom et al., 2004, Murray et al., 2003). Due to limitations in the MCS data this could not be explored within the current analysis yet still offers a likely explanation of why postnatal depression appears to be so important for child development.

Despite the finding that treatment may provide some benefit for children up to age five, the effect was not maintained at age seven where both groups of children with depressed mothers were found to be equally as likely to score high on the SDQ.

In the present analysis the differences in effects between children of undiagnosed and diagnosed mothers are shown to decrease over time, with both groups of children being equally likely to score above threshold for behaviour problems by age 7. In observing the changes in mean SDQ scores over time it was found that scores for the diagnosed group increase by 7 tenths of a point from age 5 to age 7, from 8.3 points to 9 points (Figure 4.8). In contrast the mean scores remain far more consistent over the same time period for the children in the undiagnosed group, increasing between age 5 and 7 by only 1 tenth of a point from 9.3 to 9.4.

One possible explanation which may account for why significant differences were not maintained at age seven is differences in the long-term trajectories of maternal mental health. As discussed in Chapter 2 there may be an accumulation effect of exposure to maternal depression whereby repeated/ prolonged exposure to maternal depression has a

greater effect on development than a single episode of depression. The increase in the mean behaviour score over time may reflect this effect. For instance, mothers with diagnosed depression may have more persistent or severe depression pathways. This may also be a potential reason why these women were more readily identified by health professionals or were more likely to identify themselves.

Subgroup analysis could be used to examine this further by mapping pathways of depression over time and exploring how different pathways affect the likelihood of a child scoring above threshold for behavioural problems.

One further possible explanation as to why the groups may converge by age 7 is school entry and education. It may be that been in the presence of other children or facing disciplinary action from teachers aids to improve behaviour. However children of both diagnosed and undiagnosed depression are both 2.3 times more likely to show signs of behavioural problems at age 7 suggesting that although being in school may be having a positive effect on behaviour a clear difference remains between these children and control children. More research would be required to investigate this notion.

It may also be that mothers in the diagnosed group have more severe depression, however, scores on the self-reported scale for these women would suggest evidence for lower levels of symptoms, as only 51% of women who were receiving treatment for depression scored above threshold on the Malaise Inventory, implying that these women on average did not have more severe levels of depression at this time point. Additionally the mean Malaise Inventory score for these women was 3.7, slightly lower than the threshold score of 4 required to be classified as 'depressed', and the mean score of the undiagnosed group (mean = 4.9 points). This could be indicative of a treatment effect in terms of symptom relief but this cannot be substantiated without information concerning the history of depression and the type and duration of treatment received. As the mean depression scores showed very little difference between groups, the increased effect on child behaviour problems associated with diagnosed depression is more likely a consequence of the increased likelihood of persistent depression as will be explored in Chapter 5.

4.4.7 Implications and further research

Previous studies have placed estimates of postnatal depression in the general population at around 13% (O'Hara and Swain, 1996). However, there have been suggestions that the actual rate may in fact be much higher (Hearn et al., 1998). This analysis shows that the proportion of women experiencing symptoms of postnatal depression to be around 17.7% but just over half (51.2%) of these women reported that they had not being diagnosed by a doctor as depressed. This supports the findings of Hearn and colleagues and implies that rates of postnatal depression may be higher than previously thought with a large proportion of cases being missed. It should however be acknowledged that the MCS oversampled for poverty and ethnic minorities, factors which may potentially increase the likelihood of developing postnatal depression, thus rates may be different in different samples. Of those classed as being in poverty, 26% were allocated to one of the two depression groups compared to 15% of the not in poverty group. These differences were slightly less notable for ethnicity whereby 22% of ethnic minorities were allocated to one of the depression groups compared to 18% of the white mothers. Nevertheless this may limit generalisability to samples that have different proportions of ethnic minorities and low socio-economic status women. However, having oversampled for these sub-groups is likely to be of some benefit given that other samples have not included such women and have been restricted to English speaking women who attended appointments (Hearn et al., 1998). Thus if higher than expected rates of undiagnosed depression are being identified in these samples, it is likely that even higher rates of undiagnosed depression are present in those populations that are more at risk for developing the condition.

The results of the analysis confirm the findings of previous research that maternal postnatal depression is likely to have little effect on cognitive development in terms of language but does have significant effects on behavioural development. The results of the analysis further show that diagnosis and treatment for postnatal depression appears to mediate the impact this condition has on child outcomes but this may not be a lasting effect. This could mean that interventions for postnatal depression should be focusing not only on treating the mothers' symptoms but should also include target parenting with specific focus on the mother-infant interaction. Further research is required to assess the effects that different treatments have on child development.

4.4.8 Implications

These results clearly demonstrate that maternal depression can have long lasting detrimental effects on behavioural development. Improving diagnostic strategies should reduce the number of false negative cases of women who are currently failing to be diagnosed. Further to this, treatments for postnatal depression need to be addressed to ensure that not only the maternal symptoms are being treated but also the effects of the depression on the child's development are being minimised.

4.4.9 Strengths and limitations

4.4.9.1 Strengths

The main strength of this analysis is that it is the first analysis that has explored the longitudinal impacts of undiagnosed maternal depression during the postnatal year.

A further strength of this work comes from the use of such a large scale nationally representative data set. The millennium cohort study provides a large, contemporary, and highly representative UK data set in which to explore the effect of maternal depression on child behavioural development over time. As mentioned previously the sampling strategy of the MCS allowed for over sampling of families from less affluent areas and areas of increased ethnic minorities.

Although missing data is limitation, a particularly strong point of this analysis is the way in which the potential effects of missing data were accounted for through the implementation of data imputation techniques. This is especially beneficial when using large longitudinal data sets like the MCS that are often subject to high levels of attrition and non-response to individual questions over the follow-up period. In this instance very little difference was found between the observed estimates and the imputed estimates allowing for greater confidence in the results.

A further important strength of this analysis is that the SDQ scores were analysed in both their raw and dichotomised form. Results of validation studies generally confirm that

children with 'high' SDQ scores have greater rates of mental disorder than children with 'low' SDQ scores (Goodman, 2001) but there is still a question of whether applying thresholds is appropriate (Goodman and Goodman 2009). Research that has validated the SDQ as a screening tool for behavioural problems has generally done so by comparing scores with other measures of behavioural problems such as The Child Behaviour Checklist (Goodman and Scott, 1999). Less research has been done to validate the SDQ scoring systems with actual clinical diagnosis of behavioural difficulties. In a survey analysis of two large scale nationally representative surveys of children between 5 and 16 years, SDQ's were completed by parents and teachers (Goodman and Goodman, 2009). Children were then assessed using the 'Development and Well-being Assessment (DAWBA)', a detailed psychiatric interview which allows for a clinical diagnosis to be made. A total of 7912 children were included in the final sample. The results showed children with higher total difficulty scores are more likely to be clinically diagnosed with a behavioural problem. The odds of having a behavioural disorder actually increased on a per point basis at a constant rate across the range (OR 1.28 (95% CI 1.27-1.29) per 1 point increase). This study showed no evidence of threshold effects for the SDQ at either high or low scores. In contrast to this, Goodman (2001) suggests that applying thresholds to the SDQ score to dichotomise between low risk and higher risk children is a valid methodology. A nationwide epidemiological sample of 10,438 British children was used whereby parents provided SDQ scores for their children between the ages of 5-15 years. These were compared again with the probability of clinical diagnosis according to the DAWBA. Children scoring above the 90th percentile were deemed as 'high risk' with those scoring below deemed as 'low risk'. Of the group of children classified as 'high risk' of behavioural problems 47.3% obtained a clinical diagnosis compared to only 5.4% of the 'low risk' group. This shows that scoring in the upper 10th percentile substantially raises the probability of diagnosed psychiatric disorders in children (OR 15.2 (95% CI 13.0-17.7)).

Due to the fact that there is such limited research in this area it is difficult to make an informed decision on which methodology of analysis is most appropriate. Consequently the present analysis analysed the results both on a point by point basis and also as a dichotomised outcome, thus being in line with both of these studies. As the results from both of these analyses revealed the same pattern of results this provides greater confidence in the findings.

4.4.9.2 Limitations

Despite the strengths of this analysis, using secondary data does impose some limitations. The Malaise Inventory was utilised as the measure of depressive symptoms, which although has shown good sensitivity and specificity at detecting depression (Rogers et al., 1999), a questionnaire more targeted towards postnatal depression such as the Edinburgh Postnatal Depression scale that specifically targets symptoms of postnatal depression would have been preferable. In addition, the reduced version of the Malaise Inventory used in the MCS has not been rigorously tested as a screening tool for postnatal depression. Despite this the reduced version of the Malaise Inventory has been used in comparative studies (Bartley et al., 2004, Mensah and Kiernan, 2010, Kiernan and Huerta, 2008), and is still likely to provide a good indication of the levels of depression mothers are experiencing. In addition to this, the diagnosed and treated group was derived from affirmative answers to two questions, firstly 'Has a doctor ever told you that you suffer from depression or serious anxiety?' and secondly 'And are you currently being treated for this?'. The MCS did not provide information on the history of the mother's depression meaning that we cannot be certain that the depressive episode identified at the nine month sweep was specific postnatal depression or more generalised depression that was ongoing. When examining diagnosed mothers scores on the Malaise Inventory, only 51% scored above threshold for depression. This could mean that there is a limitation with the screening measure, e.g. it has low sensitivity, or mothers in this group are recovering. However, we would expect women who are receiving treatment for depression to have different levels of symptoms, particularly over time when women are responding to treatment and symptoms are improving.

These questions also posed problems for allocation to the 'undiagnosed' group. Of the 2693 women that said they had been diagnosed with depression or serious anxiety at some time, 1736 (64%) were not receiving treatment at the time of the survey. These were then allocated to the undiagnosed or non-depressed groups according to their score on the Malaise Inventory. As a consequence of this, the group considered to be 'undiagnosed' may be contaminated by women who have recently been diagnosed with depression but who are awaiting or have refused treatment but there is no way of discerning this. Only 364 (21%) women reported having ever been diagnosed with depression but who were not

receiving treatment were subsequently assigned to the undiagnosed group thus the number of women who this applies to is likely to be small.

Women who were recently diagnosed with depression but had refused treatment may be systematically different from other women in the undiagnosed group. However, there is no way to address this within the MCS sample. This is thus acknowledged as a limitation of the data and is a potential avenue for future research.

Despite this, the present analysis is still able to give us an indication of the impact of depression during the first postnatal year on child behaviour, regardless of its origin. In relation to this, a further limitation of the MCS is the lack of insight into the specific type of treatment the women were receiving, which could take any number of different forms, meaning that comments cannot be made on the efficacy of individual treatments.

4.4.10 Further research

This analysis has highlighted several areas where further research is required. Firstly, very little documentation was found to be available concerning the ideal analytical strategy for analysing the SDQ data, with some research suggesting it should be analysed continuously and some dichotomously as noted above. Further research should thus be conducted to investigate further the validity of applying thresholds to the continuous SDQ data, dichotomising children into high and low risk groups.

In addition to this the analysis should be replicated on more longitudinal data to observe whether the effects of postnatal depression on child behaviour are persistent over time or whether they dissipate. The analysis should also be replicated on other data sources such as the Born in Bradford Cohort study (www.borninbradford.nhs.uk) to test the validity of the results on a different population with a different demographic profile (higher ethnic minority content).

As the current analysis could not establish the differential effects on child outcomes of different types of treatment, such as CBT versus home-based counselling or drug-therapy, as discussed in Chapter 1, future research should aim to identify these differences thus

providing important implications for policy and treatment development and recommendations.

One of the most important avenues of research identified in this analysis is the exploration of long-term trajectories of maternal depression and how these affect children. It is possible that children of mothers with persistent depression, either prolonged or recurring, may be more substantially affected than those whose mothers have only one episode of depression. Sub-group analysis differentiating between the potential longitudinal pathways of depression is undertaken in Chapter 5 to explore this.

4.5 Conclusions

Depression in the first postnatal year is an underdiagnosed condition with a large proportion of mothers experiencing depressive symptoms without identification or intervention. Children of depressed mothers, diagnosed or undiagnosed, are at greater risk of developing behavioural problems compared to children of non-depressed mothers. Diagnosing and treating depression in the first postnatal year appears to hold some short-term benefits for children. These findings suggest a need for better identification of depressed women and also broader recognition of the impact depression has on children. Treatments that have proven successful in treating maternal symptoms such as CBT and counselling therapies, should incorporate measures to help prevent the development of behaviour problems in children, perhaps through more family based interventions or education in parenting skills. More research is required to investigate what treatment methods have the best outcomes for both the mother and the child and the effects of differential maternal mental health trajectories over time.

Chapter 5: Persistency of maternal depression

5.1 Introduction

In general, depression as a whole is recognised as a heterogeneous disorder with a highly variable developmental course, an inconsistent response to treatment, and often no established mechanism (Belmaker and Agam, 2008). Equally, postnatal depression is highly variable in course, with some women spontaneously recovering in a short time whilst others go on to suffer prolonged or recurrent episodes of depression, is inconsistent in response to treatment, largely driven by treatment preferences and ability to maintain appropriate treatment such as CBT with a small child, and has no particular established mechanism. The only real difference between postnatal depression and general depression is the specific onset following child birth. Depression is often a long-term condition, sometimes life-long for some individuals (Blier et al., 2007). Studies examining the longitudinal course of depression have shown varying results. For instance, one general population study showed that around half of those having an episode of depression, treated or untreated, recover and remain well, whilst the other half will experience at least one subsequent episode, with 15% experiencing chronic recurrence (Eaton et al., 2008). Conversely, other studies have shown that 80% of those experiencing a depressive episode will experience at least one subsequent episode (Fava et al., 2006).

The recurrence of symptoms is thought to be more likely if treatment has not been successful at fully resolving symptoms, or if the cause of depression is unresolved, for example if depression is caused by persistent poverty. Evidence from randomised controlled trials presented in a systematic review indicates that continuing drug therapy (antidepressants) after apparent recovery can reduce the likelihood of relapse by up to 70% (41% relapse on placebo compared to 18% relapse on antidepressant) (Geddes, 2003).

Like general depression, maternal depression with a specific postnatal onset can also mark the beginning of a chronic condition, with women who experience one episode of depression being around six times more likely to have future recurrent depressive symptoms than those without depression (Josefsson et al., 2007). This means postnatal

depression constitutes as a risk factor for future mental health problems thus long-term mental health trajectories are important in examining child outcomes in relation to postnatal depression.

In Chapter 4 the effect of undiagnosed maternal depression during the first postnatal year was examined in relation to child cognitive and behavioural development in a large nationally representative UK sample drawn from the Millennium cohort study. The effect of undiagnosed depression was compared with that of diagnosed depression to explore the impact of identifying and treating maternal symptoms and to assess whether this was beneficial to child outcomes. Although significant differences were found in behavioural outcomes between children of diagnosed and undiagnosed mothers up to age 5, these were not maintained to age 7 where both groups of children scored similar scores on the SDQ and were equally likely to score above threshold for behavioural problems.

The findings of Chapter 4 provide some support for the critical periods model described in Chapter 2 by showing that exposure to depression during the first year of life has lasting effects on child development. This is indicated by the consistently significant differences observed between children of depressed mothers and non-depressed mothers. However, the fact that differences between children of mothers according to whether depression was diagnosed and treated or self-reported was not sustained over time suggests that there may be something else affecting child outcomes by age 7.

The accumulation model builds on the critical periods model, suggesting that although critical periods are important for development, effects of exposure to negative stimuli during these periods can still be modified by future events, particularly by repeated exposure to the same negative event or the accumulation of multiple negative events such as the combined effects of being raised in poverty by parents with low levels of education, both of which are independently associated with child outcomes (Chaimay et al., 2006, Peacock et al., 2013, D'Onise et al., 2010). Given that maternal depression has such large effects on children's development and is known to often be a chronic condition, either occurring over extended periods of time or presenting as recurring episodes it is reasonable to suggest that there may be a cumulative effect of exposure to prolonged or recurrent depression.

5.1.1 Persistent maternal depression across the early years of childhood

The majority of what we know about maternal postnatal depression comes from studies conducted in the immediate first postnatal year, with little consideration given to how these depressive symptoms can develop or persist over time. Depressive symptoms in women with young children have shown a high level of consistency throughout their children's early childhood, with women who report symptoms of depression during the first postnatal year being highly likely to continue to report such symptoms over time (Horwitz et al., 2007, Mclennan et al., 2001). The accumulation model drawn from the life course approach to health inequalities emphasises the importance of risks to health throughout the life course focusing upon how circumstances and experiences accumulate over the life course and create inequalities whilst accounting for developmental periods during which the impact of negative circumstances may be greater (Galobardes et al., 2004). Given what we know about the lasting impact of depression in the first postnatal year on child outcomes, the accumulation model would suggest that repeated or prolonged encounters with maternal depression could have more extensive and damaging impacts than an isolated exposure, thus making persistent maternal depression an important issue.

5.1.2 Consequences of persistent depression

Although a large body of literature documents the association between the presence of maternal postnatal depression and child outcomes, relatively less research has examined how the course, timing and persistency of the mother's depression are related to outcomes for the children of depressed mothers.

The literature that does exist converges on the idea that the more persistent the depression the more child outcomes are affected. One study conforming to this idea assessed maternal depressive symptoms at four time points and child outcomes at age five years in a large community sample (Brennan et al., 2000). It was found that mothers with more persistent and severe depression as measured by self-report using Delusions-Symptoms-States Inventory (Bedford and Foulds, 1978) had children with more severe behavioural problems according to maternal reports on the 'Child Behaviour Check List' (CBCL), especially when both severity and persistency of symptoms were enhanced. Similar results were also found in a more recent study by Ashman et al. (2008). Again, the relationship between the

longitudinal course of maternal depression and children's psychophysiology and behavioural development were examined, this time over a slightly longer period of time up to age six and a half years. Growth mixture modelling was used to identify classes of depressed mothers based on the longitudinal course of the mother's depression. Unlike Brennan and colleagues, Ashman et al. examined the effects of persistent depression on specific types of child behaviour. Children of mothers with persistent depression were found to have increased levels of overall externalising behaviour problems especially hyperactivity and attention problems compared to children of non-depressed mothers.

Previous work utilising the MCS has examined how depression at the first two time points affects developmental outcomes (Kiernan and Mensah, 2009). Mothers were classified as depressed at the 9 month data sweep if they reported that they had felt low or sad for a period of two weeks or more since the child's birth. A similar self-reporting measure was used when the child was 3 years old. When depression was only reported at sweep 1 children were 1.5 times more likely than controls to show behavioural problems, but, when depression was reported at both sweeps children were 3.74 times more likely than children of non-depressed mothers to show behavioural problems showing a clear effect of persistent depression. Such effects were significant even after accounting for potential confounding variables.

Although these studies clearly show that maternal depression that persists or recurs over time stronger associations with children's behavioural development than postnatal depression alone, little is known about how changes in maternal mood over time affect children on a longer term scale. Further to this, no previous work was identified that has explored the effects of long-term changes in maternal mood following depression in the first postnatal year according to whether depression was diagnosed and treated or unidentified.

5.1.3 Aims

This chapter aims to examine the longitudinal course of maternal depression according to diagnosis and treatment and to explore the effects these different pathways have on child behavioural outcomes up to age 7 using a sample from the Millennium Cohort Study.

This study investigates the potential pathways of maternal depression over time taking into consideration whether depression was diagnosed and treated or not in the postnatal year (sweep 1, 9 months). The relationship between different pathways of depression over time and child behavioural outcomes at age seven is explored.

5.1.4 Research questions

The following questions are addressed;

- What are the potential pathways maternal depression can take over time?
- How does diagnosis and treatment impact future depression pathways?
  - Are those who are undiagnosed and untreated more likely to remain depressed for longer?
- How do the different pathways of depression affect child outcomes?
- Is there evidence for a cumulative effect of exposure to maternal depression?
- Does depression only occurring in the first postnatal year have lasting effects?
- Is the first postnatal year a 'critical period' to intervene?

5.2 Methods

5.2.1 Sample

Data for this study was again drawn from the first four sweeps of the Millennium Cohort Study (MSC). Mothers must have provided information on mental health at each of the four sweeps and information on child behaviour at age 7. Based on these criteria a total of 6240 mothers were used in this analysis. This requirement resulted in a reduction in the sample size from that used in Chapter 4 in the multi-level model as mental health data was missing from some mothers at the 3, 5 and 7 year follow-up sweeps. The demographic attributes thus differ slightly in this sample compared to that used in the analysis presented in Chapter 4 (Table 5.1). As can be seen in Table 5.1 there are some differences in the two samples. The sub-sample to be used in the present analysis as a whole contains mothers that are more educated, reflected in lower proportions of mothers reporting holding no qualifications across all three groups, more likely to be white, and less likely to be

considered in poverty thus there is some potential for reduction in the generalisability from the previous sample. However, weighting, as discussed in Chapter 4, section 4.2.2, was again applied to the current sample in the analysis which alters the contribution of cases in statistical models improving the generalisability.

Table 5.1 Participant demographic characteristics according to maternal mental health group for the sub-group of families used in the persistency analysis

	Diagnosed N= 668	Undiagnosed N= 622	Non-depressed N= 4950
Maternal characteristics			
Age at CM birth (years), mean	28.78	28.07	29.88
Qualification Level			
Level 1	8.53	9.81	5.74
Level 2	35.18	30.06	27.71
Level 3	16.92	16.40	16.05
Level 4	24.55	26.05	39.39
Level 5	1.95	3.54	4.93
Overseas Only	1.80	3.05	1.09
None	11.08	11.09	5.09
Lone parent			
Yes	17.66	18.81	9.07
No	82.34	81.19	90.93
In poverty			
Yes	32.97	33.79	17.81
No	67.03	66.21	82.19
First live birth			
Yes	36.45	39.77	43.09
No	63.55	60.23	56.91
In employment			
Yes	44.30	49.11	59.86
No	55.69	50.89	40.14
Ethnicity			
White	96.25	90.35	95.30
Mixed	0.15	0.64	0.61
Indian	1.20	2.09	1.22
Pakistani/ Bangladeshi	1.50	3.54	0.89
Black/ Black British	0.75	2.41	0.89
Other	0.15	0.96	0.57
Child characteristics			
Child age (months), mean	9.15	9.16	9.19
Gender			
Male	47.90	52.73	50.12
Female	52.10	47.27	49.88
Ethnicity			
White	94.15	89.07	94.03
Mixed	2.25	2.25	2.17
Indian	1.20	1.93	1.19
Pakistani/ Bangladeshi	1.35	3.54	0.87
Black/ Black British	0.90	2.73	1.40
Other	0.15	0.48	0.34

5.2.2 Groupings

For this analysis mothers are initially classified into one of three groups at sweep 1 (age 9 months):

1. Mothers who were currently receiving treatment for depression at the time of the sweep.
2. Mothers with high levels of self-reported depressive symptoms but who were not currently receiving treatment.
3. Control mothers who were not identified as depressed on any measure.

And one of two groups at sweeps 2-4 (ages 3, 5 and 7 years):

1. Mothers who are depressed (diagnosed or undiagnosed)
2. Mothers who are not identified as depressed on any measure

5.2.3 Covariates

The present analysis will utilise the same covariates as those used in the analysis in Chapter 4 as identified as important for child outcomes in Chapter 3. These include: household poverty, mothers' level of education, whether the mother is a single parent or not, ethnicity, mothers' age at cohort members birth, whether the cohort member is the mother's first child, whether the mother is employed, the child's gender and the child's age at time of follow up.

5.2.4 Weighting

As in Chapter 4 weighting strategies will need to be employed to take account of the initial sampling design as well as non-response in the recruitment of the original sample and sample attrition over the follow up period up to age seven. This allows for a more representative and generalizable representation of the UK to be drawn.

5.2.5 Measures

5.2.5.1 Self-reported mental health scales

At the nine month survey mothers completed a nine-item reduced Malaise Inventory (Rutter et al., 1970) whereby a score of at least four points was used to classify high levels of depression. As mentioned in Chapter 4, the Malaise Inventory includes emotional disturbance as well as somatic symptoms, and has been widely used as a measure of distress and depression levels (Grant et al. 1990; McGee et al. 1986). The inventory has been shown to have good reliability and validity (Grant et al. 1990).

At the following surveys mothers completed the Kessler 6 scale (Kessler et al. 2002, Kessler et al. 2003) via a computerised self-report whereby a score of at least seven points is indicative of depression.

The diagnosed and in treatment mothers were identified at each sweep through answers to the following interview questions;

1. 'Has a doctor ever told you that you suffer from depression or serious anxiety?'
2. 'And are you currently being treated for this?'

5.2.5.2 Child outcome measures

i. Behavioural development

Behavioural adjustment was assessed at age three used the Strengths and Difficulties Questionnaire (SDQ) (Goodman, 1997) which forms a 25 item behavioural screening questionnaire on five different dimensions of children's behaviour including; conduct problems inattention-hyperactivity, emotional symptoms, peer problems and pro-social behaviour. Each attribute was rated by the mother using a scale from 0-2 (not true, somewhat true, or certainly true). Responses were summed to provide a total score for each dimension. As in Chapter 4 both the continuous total score and the dichotomised threshold based score will be used in the analysis.

ii. Cognitive development

As the results in Chapter 4 showed that diagnosis of maternal depression does not impact child cognitive development this will not be explored as an outcome in the present analysis as it would not add weight to the decision model for screening.

5.2.6 Statistical analysis

Data were collected at each sweep regarding the mothers' mental health allowing mothers to be categorised as depressed and diagnosed (receiving treatment), depressed and undiagnosed (not receiving treatment) and not depressed at each time point. Pathways of maternal depression were developed by tracking mothers' depression status over time.

Only mothers who were categorised as not depressed at all four sweeps were included in the reference group as the aim of the analysis is to explore the importance of diagnosing maternal depression during the first postnatal year, how this affects trajectories of maternal mental health, and how these trajectories impact on child outcomes. Though maternal depression arising outside the first postnatal year is important, it is irrelevant for the current question.

Initially the three categories (diagnosed/ undiagnosed/ non-depressed) were used at each sweep to determine the pathways however this resulted in an unfeasible number of pathways reduce the number of women in each group by too much to allow for statistical analysis.

Methods of compressing the pathways were thus explored. A sequential series of four models were developed taking an informed decomposition modelling approach. Each model was grounded on knowledge drawn from theoretical framework outlined in Chapter 2, namely the life course approach. This allowed for an intuitive approach to pathway decomposition. Continuous outcome measures were modelled using linear regression which dichotomous measures were modelled using logistic regression.

Model 1: Aimed to utilise all available data concerning maternal mental health at each of the three further time points stemming from the initial grouping at time one (9 months).

This mapped mothers according to the three groups, diagnosed, undiagnosed and non-depressed, at each of the four time points creating 55 pathways. Due to the limited number of mothers following certain pathways within the model it was deemed inappropriate to conduct statistical analysis using this model, thus alternatives were explored.

Model 2: Aimed to condense the 55 pathways to allow for statistical and more adequately reflect the actual pathways followed by women. As depression during the first postnatal year and screening is the focal area of interest for this thesis, diagnostic status is only truly important during this period, as this is the time women would receive the screening. Consequently, it was decided that depression at subsequent sweeps would be categorised only as 'depressed' or 'not-depressed', combining mothers who were identified as depressed on either measure in order to reduce the number of pathways. At sweep 1 mothers remained categorised according to the original three definitions in order to assess whether diagnosing and treating at this time point impacted upon child outcomes after future incidences/persistent depression was considered. This resulted in the derivation of 17 pathways. This model was assessed statistically using regression modelling. As the outcome of interest was child behavioural development at age 7 only, a multilevel model to account for the effects of repeated measures over time was not required for this analysis, thus standard linear logistic regression modelling were employed for the continuous outcome and the dichotomous outcome respectively.

Model 3: Model 3 aimed to reduce the pathways further. As shown in Table 5.2 a large proportion of women from both the diagnosed and undiagnosed groups follow the persistent depression (depression at all 4 time points) or the recovery pathway (depression only at sweep 1).

These pathways account for 53% of undiagnosed women (18.6% persistent, 34.4% recovered) and 57% of diagnosed women (29.6% persistent, 17.4% recovered). The remainder of the women follow an intermittent pathway with depression reported on at least one further follow up sweep.

Because the majority of women in each of the diagnosed and undiagnosed groups follow 2 out of 8 potential pathways per group, group numbers for the remaining pathways are low which reduces statistical power.

Model 3 compresses the 12 intermittent pathways (6 for diagnosed women and 6 for undiagnosed women) into 2 'intermittent' groups. This resulted in the derivation of 7 pathways. As is model 2, this model was assessed statistically using regression modelling. As the outcome of interest was child behavioural development at age 7 only, a multilevel model to account for the effects of repeated measures over time was not required for this analysis, thus

Model 4: Aimed to validate the results of model 3. When categorical variables are condensed, the clarity and richness of the data can be reduced. As an incremental effect was found in model 3 whereby the odds of children presenting with behavioural problems on the SDQ scale at age 7 increased in line with the persistency of the mothers depression, it is hypothesised that there may be a cumulative effect of exposure to episodes of depression.

In Model 3 an 'intermittent' category was derived to compress pathways. One problem with this is that mothers who reported depression at 2 time points are grouped with mothers who reported depression at 3 time points. According to the principles of the life course approach discussed in Chapter 2, the more frequent the exposure to negative stimuli the greater the effects may be, thus it is important to consider the number of exposure time points. Model 4 thus sought to expand the 'intermittent' category back out by categorising mothers according to the number of exposures reported. These could be in any order after the first sweep during the post natal period (9 months).

Women were thus separated into those identified as having 1, 2, 3, or 4 episodes of depression, 1 of which must be during the first postnatal year at sweep 1. This resulted in 9 pathways. As with the previous two models standard linear logistic regression modelling were employed for the continuous outcome and the dichotomous outcome respectively.

Each of the three analytical models were conducted firstly on a univariate basis and secondly on a multivariate basis with covariates (as identified in Chapter 3). This approach was used to give an indication of the effects adjusting for the covariates identified as important has on the model, compared to examining the effect of maternal depression alone.

5.2.7 Missing data

Multiple imputation analysis was carried out in Chapter 4 to assess the impact of missing data within the MCS sample. As this imputation analysis showed that missing data has very little effect in this sample with regards to the variable of interest it will not be repeated on the present analysis.

5.3 Results

5.3.1 Model 1

In the first instance maternal depression at each time point was classified according to the type of depression, diagnosed or undiagnosed. This generated 55 pathways (Figure 5.1). Figure 5.1 is colour coded to show the changes in maternal mental health states over time, colours used for each of the mental health states are presented in Box 5.1. The number of women following each of the pathways is presented alongside the number of children behavioural outcome data is available for per group.

Box 5.1 Colour coded key for the regression models

Pathways	Sweep 1	Sweep 2	Sweep 3	Sweep 4	N children	N mothers
1	green	green	green	green	4950	5,604
2	red	green	green	green	183	209
3	red	green	green	red	26	28
4	red	green	green	yellow	14	16
5	red	green	red	green	22	28
6	red	green	red	red	30	37
7	red	green	red	yellow	5	5
8	red	green	yellow	green	9	13
9	red	green	yellow	red	4	10
10	red	green	yellow	yellow	6	6
11	red	red	green	green	51	59
12	red	red	green	red	21	26
13	red	red	green	yellow	12	13
14	red	red	red	green	23	30
15	red	red	red	red	119	160
16	red	red	red	yellow	13	14
17	red	red	yellow	green	17	19
18	red	red	yellow	red	15	17
19	red	red	yellow	yellow	11	13
20	red	yellow	green	green	22	23
21	red	yellow	green	red	5	5
22	red	yellow	green	yellow	7	11
23	red	yellow	red	green	4	4
24	red	yellow	red	red	7	11
25	red	yellow	red	yellow	8	10
26	red	yellow	yellow	green	9	11
27	red	yellow	yellow	red	8	11
28	red	yellow	yellow	yellow	17	23
29	yellow	green	green	green	214	254
30	yellow	green	green	red	20	24
31	yellow	green	green	yellow	35	42
32	yellow	green	red	green	8	12
33	yellow	green	red	red	9	14
34	yellow	green	red	yellow	4	4
35	yellow	green	yellow	green	21	26
36	yellow	green	yellow	red	5	8
37	yellow	green	yellow	yellow	18	27
38	yellow	red	green	green	26	29
39	yellow	red	green	red	1	2
40	yellow	red	green	yellow	8	8
41	yellow	red	red	green	6	7
42	yellow	red	red	red	26	34
43	yellow	red	red	yellow	3	6
44	yellow	red	yellow	green	8	8
45	yellow	red	yellow	red	6	7
46	yellow	red	yellow	yellow	8	9
47	yellow	yellow	green	green	62	75
48	yellow	yellow	green	red	5	9
49	yellow	yellow	green	yellow	14	18
50	yellow	yellow	red	green	8	8
51	yellow	yellow	red	red	7	9
52	yellow	yellow	red	yellow	5	8
53	yellow	yellow	yellow	red	16	20
54	yellow	yellow	yellow	green	34	39
55	yellow	yellow	yellow	yellow	45	56

Figure 5.1. The range of potential pathways when diagnosis status is considered at each time point

As can be seen in Figure 5.1, mapping maternal depression into 55 pathways resulted in vast reductions to the number of mother-child pairs within each group, with many pathways having figures less than 10. As a consequence of this further reduction was required to ensure groups were statistically comparable and were more reflective of the mental health trajectories mothers actually follow.

5.3.2 Model 2

In the context of the current research and the particular area of interest being depression in the first postnatal year, diagnostic status is only truly important during this period, as this is the time women would receive the screening. Consequently, it was decided that depression at subsequent sweeps would be categorised only as 'depressed' or 'not-depressed', combining mothers who were identified as depressed on either measure in order to reduce the number of pathways. At sweep 1 mothers remained categorised according to the original three definitions in order to assess whether diagnosing and treating at this time point impacted upon child outcomes after future incidences/persistent depression was considered.

This resulted in the derivation of 17 pathways depicted in Figure 5.2 and described in Table 5.1. Pathway17 refers to mothers who were not depressed at any of the 4 time points on either measure of depression.
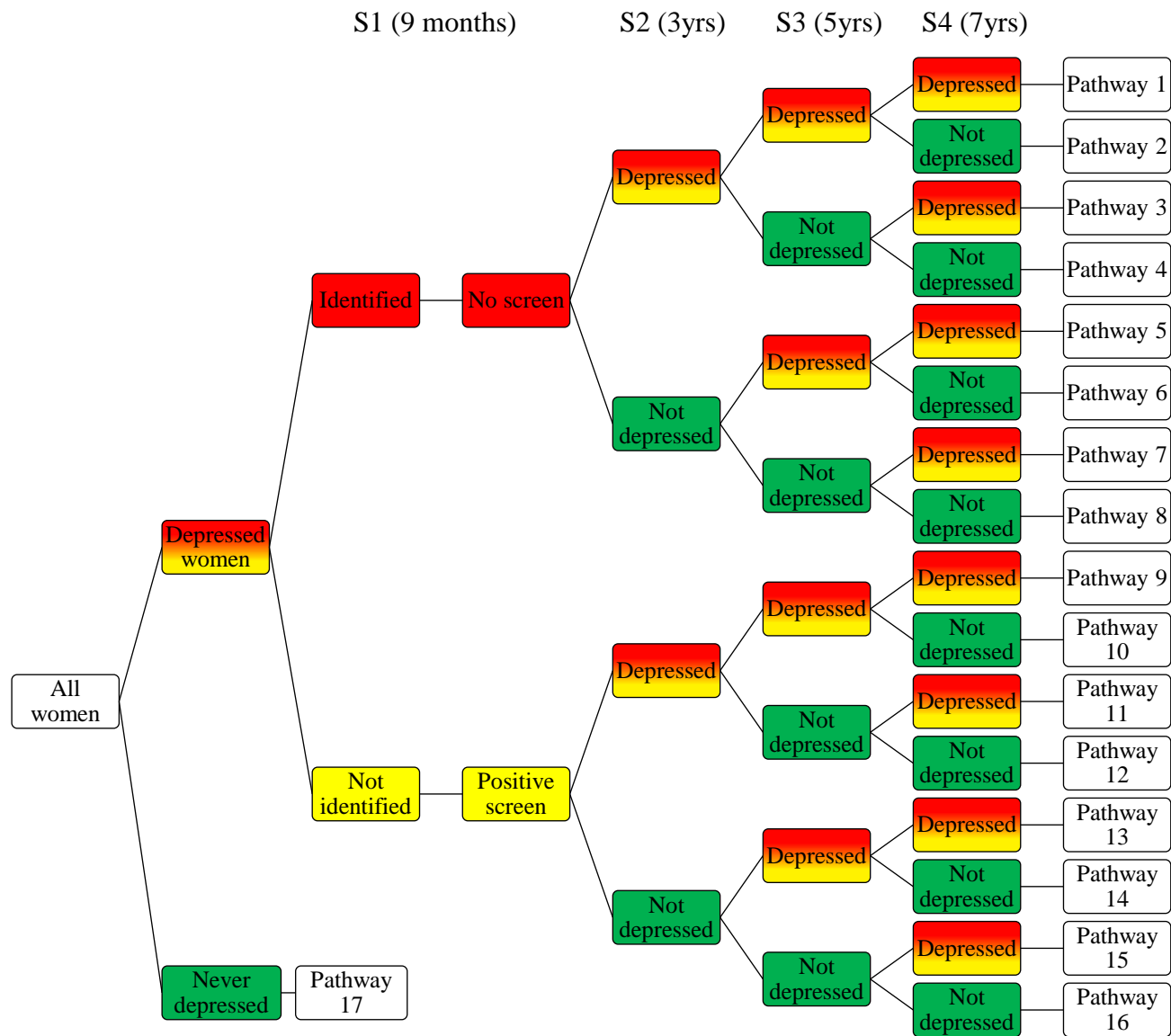
Figure 5.2. Decision model of depression over time mapping 17 pathways

As can be seen in Table 5.2 the proportions of mothers vary between pathways with some pathways being more frequently followed than others. For example, for those mothers who were diagnosed at sweep 1, more go on to report depression at all three further sweeps (n=198, 29.6% of diagnosed women) than to report recovery at all further sweeps (n=183, 27.4% of diagnosed women).

Large differences can also be seen in the proportions of children scoring above threshold for behavioural difficulties between the different groups with higher proportions of behavioural problems occurring in line with more frequent/persistent depression.

Table 5.2. Proportion of mothers following each of the 17 pathways and associated proportions of child behaviour problems at age 7

Pathway	Code	N mother-infant pairs	N Children with behaviour problems	% Children with behaviour problems
1	$D^D DDD$	198	60	30.3
2	$D^D DDN$	53	11	20.8
3	$D^D DND$	45	8	17.8
4	$D^D DNN$	73	8	11
5	$D^D NDD$	45	12	26.7
6	$D^D NDN$	31	4	12.9
7	$D^D NND$	40	10	25
8	$D^D NNN$	183	21	11.5
9	$D^U DDD$	116	53	45.7
10	$D^U DDN$	56	14	25
11	$D^U DND$	28	9	32.1
12	$D^U DNN$	88	5	5.7
13	$D^U NDD$	36	13	36.1
14	$D^U NDN$	29	6	20.7
15	$D^U NND$	55	14	25.5
16	$D^U NNN$	214	19	8.9
17	NNNN	4950	287	5.8

Code: D^D – diagnosed depression, D^U – undiagnosed depression, D – depressed (either D^D or D^U), N – not depressed

The potential pathways of depression following initial diagnosis are the same for each of the two depression groups. This means that comparable pathways can be drawn whereby the only difference is diagnosis at sweep 1 (Table 5.3).

Table 5.3. Comparable pathways of depression according to diagnosis from the 17 pathways model

| | Diagnosed | | Undiagnosed | | |
Pathway	% Mothers	% Children with behaviour problems	% Mothers	% Children with behaviour problems	Difference in behaviour problems (U-D)
DDDD	29.64	30.30	18.65	45.69	15.39
DDDN	7.93	20.75	9.00	25.00	4.25
DDND	6.74	17.78	4.50	32.14	14.36
DDNN	10.93	10.96	14.15	5.68	-5.28
DNDD	6.74	26.67	5.79	36.11	10.56
DNDN	4.64	12.90	4.66	20.69	7.79
DNND	5.99	25.00	8.84	25.45	0.45
DNNN	27.4	11.48	34.41	8.88	-2.60

D – depressed (either D^D or D^U), N – not depressed

The results presented in Table 5.3 show that in the majority of pathways a greater proportion of children of undiagnosed mothers compared to diagnosed mothers develop behaviour problems. However, when mothers are depressed at only sweep 1, or only sweep 1 and sweep 2, a greater proportion of children in the diagnosed group develop behaviour problems, though the differences between groups are fairly small, 2.6% and 5.28% respectively. The most notable differences occur when mothers are depressed at all 4 sweeps or have a recovery period at sweep 3 only. In both cases around 15% more children of undiagnosed women compared to children of diagnosed women develop behaviour problems. With the exception of children in the undiagnosed group where mothers were depressed at sweeps 1 and 2, all pathways of depression lead to a higher proportion of children developing behaviour problems at age 7 compared to controls as can be seen in Figure 5.3.
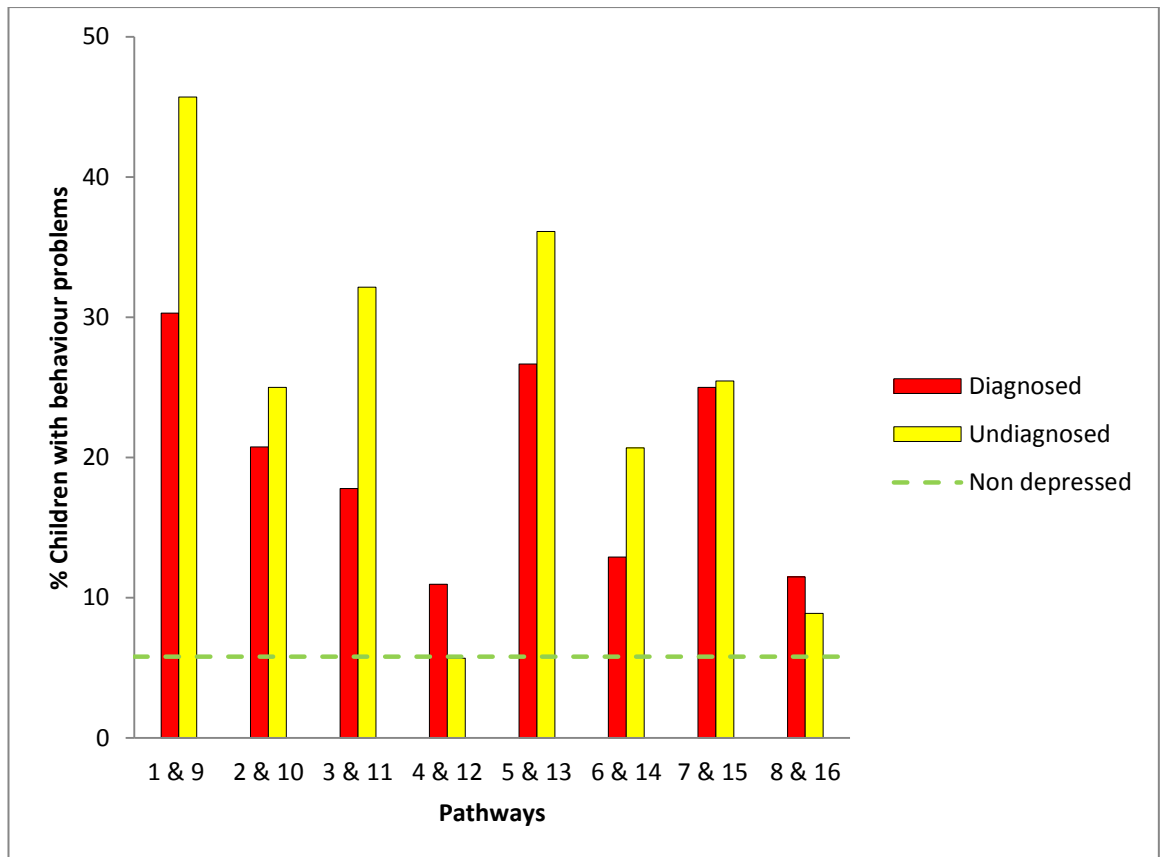
Figure 5.3. Proportions of children scoring above threshold for behavioural problems on the SDQ at age 7 according to the 17 pathway model

In order to test these associations statistically, univariate and multivariate regressions were carried out comparing each of the individual pathways to the non-depressed group and the comparable pathways to each other. These were conducted on both the continuous score and the dichotomised score as in Chapter 4 to ensure the full scope of data was captured.

Table 5.4. The effect of maternal depression pathways on child total SDQ scores at age 7: Comparison to children of non-depressed mothers

Pathway	Univariate			Multivariate		
	Coef	95% CI	P	Coef	95% CI	P
$D^D DDD$	0.83	0.72-0.94	<.0001	0.64	0.55-0.75	<.0001
$D^D DDN$	0.51	0.36-0.68	<.0001	0.51	0.34-0.69	<.0001
$D^D DND$	0.41	0.28-0.56	<.0001	0.33	0.20-0.47	<.0001
$D^D DNN$	0.11	0.00-0.24	0.042	0.04	-0.07-0.15	0.488
$D^D NDD$	0.85	0.68-1.04	<.0001	0.69	0.53-0.86	<.0001
$D^D NDN$	0.33	0.14-0.56	<.0001	0.31	0.12-0.53	0.001
$D^D NND$	0.88	0.70-1.07	<.0001	0.80	0.63-0.98	<.0001
$D^D NNN$	0.17	0.09-0.24	<.0001	0.10	0.03-0.18	0.004
$D^U DDD$	1.09	0.96-1.24	<.0001	0.80	0.68-0.92	<.0001
$D^U DDN$	0.78	0.67-0.90	<.0001	0.53	0.44-0.63	<.0001
$D^U DND$	1.01	0.80-1.26	<.0001	0.82	0.63-1.05	<.0001
$D^U DNN$	0.35	0.26-0.45	<.0001	0.24	0.15-0.33	<.0001
$D^U NDD$	0.83	0.59-1.10	<.0001	0.66	0.45-0.89	<.0001
$D^U NDN$	0.39	0.19-0.62	<.0001	0.39	0.18-0.63	<.0001
$D^U NND$	0.56	0.38-0.76	<.0001	0.49	0.33-0.67	<.0001
$D^U NNN$	0.23	0.16-0.31	<.0001	0.19	0.13-0.26	<.0001

Code: D^D – diagnosed depression, D^U – undiagnosed depression, D – depressed (either D^D or D^U), N – not depressed

Table 5.5. The effect of maternal depression pathways on children's odds of behavioural problems at age 7: Comparison to children of non-depressed mothers

Pathway	Univariate			Multivariate		
	OR	95% CI	P	OR	95% CI	P
$D^D DDD$	8.08	6.66-9.81	<.0001	6.37	5.20-7.81	<.0001
$D^D DDN$	5.16	3.42-7.78	<.0001	5.14	3.26-8.11	<.0001
$D^D DND$	1.93	1.27-2.94	0.002	1.60	1.04-2.45	0.033
$D^D DNN$	1.67	1.10-2.55	0.016	1.28	0.79-2.06	0.317
$D^D NDD$	5.88	4.04-8.55	<.0001	4.53	3.04-6.74	<.0001
$D^D NDN$	2.36	1.32-4.23	0.004	2.52	1.32-4.79	0.005
$D^D NND$	5.88	3.95-8.77	<.0001	5.68	3.71-8.68	<.0001
$D^D NNN$	2.23	1.70-2.93	<.0001	1.94	1.45-2.60	<.0001
$D^U DDD$	11.28	9.00-14.15	<.0001	7.60	5.93-9.74	<.0001
$D^U DDN$	4.88	3.43-6.93	<.0001	2.89	1.92-4.36	<.0001
$D^U DND$	8.65	5.34-14.01	<.0001	6.52	3.74-11.34	<.0001
$D^U DNN$	1.45	0.86-2.44	0.164	1.08	0.64-1.84	0.769
$D^U NDD$	8.65	5.77-12.99	<.0001	6.63	4.32-10.19	<.0001
$D^U NDN$	5.12	3.11-8.43	<.0001	5.63	3.23-9.80	<.0001
$D^U NND$	5.00	3.49-7.17	<.0001	4.79	3.33-6.88	<.0001
$D^U NNN$	1.85	1.39-2.44	<.0001	1.59	1.19-2.11	0.001

Code: D^D – diagnosed depression, D^U – undiagnosed depression, D – depressed (either D^D or D^U), N – not depressed

As can be seen in Tables 5.4 and 5.5, children whose mothers had depression at least during the first postnatal year (diagnosed or undiagnosed) generally scored higher scores on the SDQ, reflecting greater levels of behavioural problems, and were more likely to score above the designated threshold for signs of behavioural difficulties.

Univariate regression carried out on the total SDQ score showed children of mothers following all depression pathways score significantly higher than children of non-depressed mothers on the SDQ. However, after accounting for potential confounders these effects were reduced and children of mothers following pathway 4 (D^DDNN) were no longer significantly different from children of non-depressed mothers.

The pattern of results shown by the logistic regression conducted using the dichotomised outcome reveal similar results.

When regressed without covariates children whose mothers followed all pathways except pathway 12 (D^UDNN) were statistically different from children of non-depressed mothers with children of depressed mothers being more likely to score over threshold for behaviour problems on the SDQ. After accounting for covariates children whose mothers followed pathway 4 also became insignificantly different from controls suggesting the differences seen in the univariate analysis were affected by confounders.

Children whose mothers followed the most persistent depression pathways were the most likely to score above threshold for behavioural problems compared to controls with children of mothers who were depressed and diagnosed at sweep 1 and were depressed at the 3 follow ups been 6.37 times more likely to score above threshold, and children of mothers who were depressed but not diagnosed at sweep 1 as well as being depressed at all 3 follow up sweeps been 7.6 times more likely to score above threshold than children of controls.

Even when depression was only present at sweep 1 children of depressed mothers were significantly more likely than controls to score above threshold for behaviour problems at age 7, with children of diagnosed mothers been almost twice as likely (OR 1.94, p<.0001) with children of undiagnosed mothers been 1.6 times more likely (p=.001).

161

In order to assess the effects of diagnosis in relation to future depression on child behaviour the comparable pathways outlined in Table 5.3 were also compared using univariate and multivariate regression as well as logistic regression.

Table 5.6. The effect of maternal depression pathways on child total SDQ scores at age 7: Comparing children of undiagnosed mothers to children of diagnosed mothers

Pathway	Univariate			Multivariate		
	Coef	95% CI	P	Coef	95% CI	P
DDDD	0.15	0.05-0.25	0.003	0.10	0.01-0.2	0.029
DDDN	0.18	0.04-0.34	0.009	0.01	-0.11-0.15	0.901
DDND	0.42	0.23-0.66	<.0001	0.37	0.18-0.6	<.0001
DDNN	0.21	0.07-0.38	0.003	0.19	0.05-0.35	0.007
DNDD	-0.01	-0.16-0.17	0.917	-0.02	-0.16-0.16	0.851
DNDN	0.04	-0.16-0.29	0.717	0.04	-0.17-0.3	0.747
DNND	-0.17	-0.29- -0.03	0.019	-0.17	-0.28- -0.03	0.017
DNNN	0.06	-0.03-0.15	0.193	0.08	-0.01-0.17	0.081

Code: D^D – diagnosed depression, D^U – undiagnosed depression, D – depressed (either D^D or D^U), N – not depressed

After accounting for potential confounders in the multivariate analysis 4 comparable pathways showed significant differences in child outcomes in terms of continuous SDQ scores.

Children of mothers who had undiagnosed depression combined with 3 further sweeps reporting depression scored 0.10 points more than children whose mothers had diagnosed depression at sweep 1 in combination with further depression at all 3 follow ups (p=.029).

Children of mothers who had undiagnosed depression at sweep 1 followed by depression at sweeps 2 and 4 but not at sweep 3 also scored significantly higher on the SDQ than comparable diagnosed mothers. This was also the case for children whose mothers were undiagnosed at sweep 1 compared to those whose mothers were diagnosed at sweep 1 when depression was additionally present at sweep 2 but at no other sweep.

In contrast to these findings which place children at a disadvantage for behavioural problems compared to children of diagnosed mothers, an inverse relationship was found when mothers were depressed at only sweeps 1 and 4. In this case, children of diagnosed

women at sweep1 scored 0.17 points less than children of undiagnosed women, a significant difference (p=.017).

Table 5.7. The effect of 16 maternal depression pathways on children's odds of behavioural problems at age 7: Comparing children of undiagnosed mothers to children of diagnosed mothers

	Univariate			Multivariate		
Pathway	OR	95% CI	P	OR	95% CI	P
DDDD	1.40	1.05-1.85	0.020	1.19	0.89-1.61	0.247
DDDN	0.95	0.56-1.61	0.84	0.56	0.31-1.03	0.062
DDND	4.48	2.38-8.43	<.0001	4.09	2.04-8.18	<.0001
DDNN	0.86	0.45-1.67	0.666	0.85	0.42-1.72	0.648
DNDD	1.47	0.85-2.53	0.163	1.46	0.83-2.6	0.192
DNDN	2.17	1.01-4.63	0.046	2.24	0.96-5.21	0.062
DNND	0.85	0.5-1.44	0.547	0.84	0.49-1.46	0.542
DNNN	0.83	0.57-1.21	0.327	0.82	0.55-1.21	0.313

Code: D^D – diagnosed depression, D^U – undiagnosed depression, D – depressed (either D^D or D^U), N – not depressed

Although some differences were found on the continuous measure these were not well reflected in the likelihoods that children would score over threshold for behaviour problems on the SDQ. Only one comparison was shown to be significant after controlling for confounders in the logistic regression. Children whose mothers had undiagnosed depression at sweep 1 in addition to depression at sweeps 2 and 4 were 4 times more likely to score above threshold for behaviour problems compared to comparable children of diagnosed mothers.

Two other comparisons almost reached statistical significance. Children of mothers who were undiagnosed at sweep 1 and were depressed additionally at sweep 3 were twice as likely as comparable children to score above threshold though this did not reach significance (p=.062).

In line with the results relating to the continuous score children of mothers depressed at sweeps 1 and 3 performed worse when the mother was diagnosed with depression at sweep 1 compared to when mothers were depressed but without diagnosis.

5.3.3 Model 3

Model 3 aimed to reduce the pathways further. As shown in Table 5.2 a large proportion of women from both the diagnosed and undiagnosed groups follow the persistent depression (depression at all 4 time points) or the recovery pathway (depression only at sweep 1).

These pathways account for 53% of undiagnosed women (18.6% persistent, 34.4% recovered) and 57% of diagnosed women (29.6% persistent, 17.4% recovered). The remainder of the women follow an intermittent pathway with depression reported on at least one further follow up sweep.

Because the majority of women in each of the diagnosed and undiagnosed groups follow 2 out of 8 potential pathways per group, group numbers for the remaining pathways are low which reduces statistical power.

Model 3 compresses the 12 intermittent pathways (6 for diagnosed women and 6 for undiagnosed women) into 2 'intermittent' groups. This resulted in the derivation of 7 pathways (Figure 5.4).
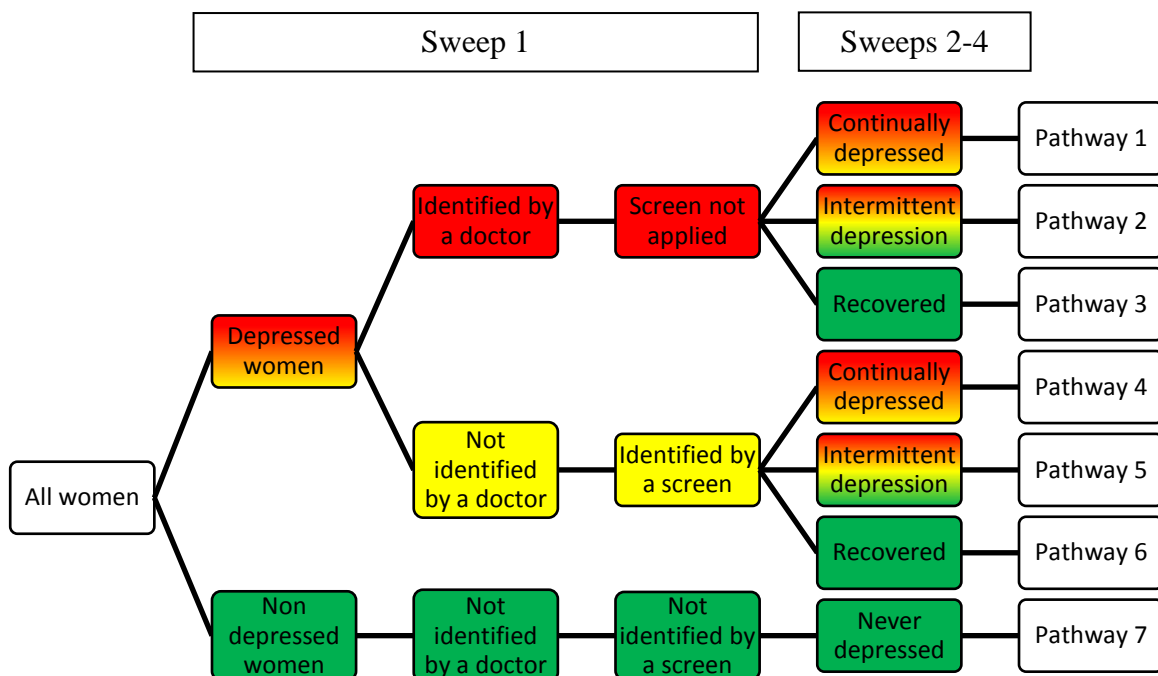
Figure 5.4. Decision model of depression over time mapping 7 pathways

The number of mother-infant pairs following each of these 7 pathways as well as the proportion of children with behavioural problems is shown in Table 5.8.

Table 5.8. Proportion of mothers following each of the 7 pathways and associated proportions of child behaviour problems at age 7

Pathway	Code	N mother-infant pairs	N Children with behaviour problems	% Children with behaviour problems
1	D^D Continuous	198	60	30.30
2	D^D Intermittent	287	53	18.47
3	D^D Recovered	183	21	11.50
4	D^U Continuous	116	53	45.70
5	D^U Intermittent	292	61	20.89
6	D^U Recovered	214	19	8.90
7	Not depressed	4950	287	5.80

Code: D^D – diagnosed depression, D^U – undiagnosed depression

As model 3 is a reduced version of model 2 some of the pathways remain the same, these are the depressed-recovered pathways and the continuous depression pathways. The only difference between model 2 and model 3 is the compression of pathways where depression status changes over time to a single pathway, thus pathway 2 is a compression of pathways 2-7 from model 2 and pathway 5 is a compression of pathways 10-15 from model 2.

As can be seen in Table 5.8, reducing the number of pathways to 7 results in much more equal numbers of women within each of the depression groups with the majority of women who were classified as depressed (depressed and diagnosed) at sweep 1 following the intermittent pathways.

The raw proportions of children scoring above threshold for behavioural problems appear to be associated with the pathway of depression their mother follows. For children of mothers who were depressed at sweep 1 (diagnosed or undiagnosed) the lowest proportions of behavioural difficulties are seen in children whose mothers were only depressed at this time point, these proportions then increase for intermittent depression and again for persistent depression measured at all 4 time points. This is true for children of mothers who were diagnosed or undiagnosed.

As in model 2 the pathways within model 3 are comparable. Table 5.9 shows these comparisons.

Table 5.9. Comparable pathways of depression according to diagnosis from the 7 pathways model

| Pathway | Diagnosed | | Undiagnosed | | Difference in behaviour problems (U-D) |
	% Mothers	% Children with behaviour problems	% Mothers	% Children with behaviour problems	
D-Cont	29.64	30.3	18.65	45.70	15.4
D-Int	42.96	18.47	46.95	20.89	2.42
D-Rec	27.40	11.5	34.4	8.90	-2.6

Code: D – depressed

The results presented in Table 5.9 show that a greater proportion of children of undiagnosed mothers score above threshold for behaviour problems on the SDQ at age 7 when mothers have continuous or intermittent depression but not if maternal depression occurs at only sweep 1. Only small differences are seen in the proportions of children scoring above threshold between intermittently depressed women and women depressed only once, with only a 2.42% difference between children of intermittent depressed women and 2.6% difference between children of only postnatally depressed women.

Much greater differences are again observed between children whose mothers were continually depressed with more than 15% more children scoring above threshold for behaviour problems when their mother had undiagnosed depression at sweep 1 compared to children of mothers who had diagnosed depression at sweep 1.

As can be seen in Figure 5.5 higher proportions of behavioural problems were found in children of mothers who followed a depression pathway compared to children of mothers who remained without depression across all 4 time points.
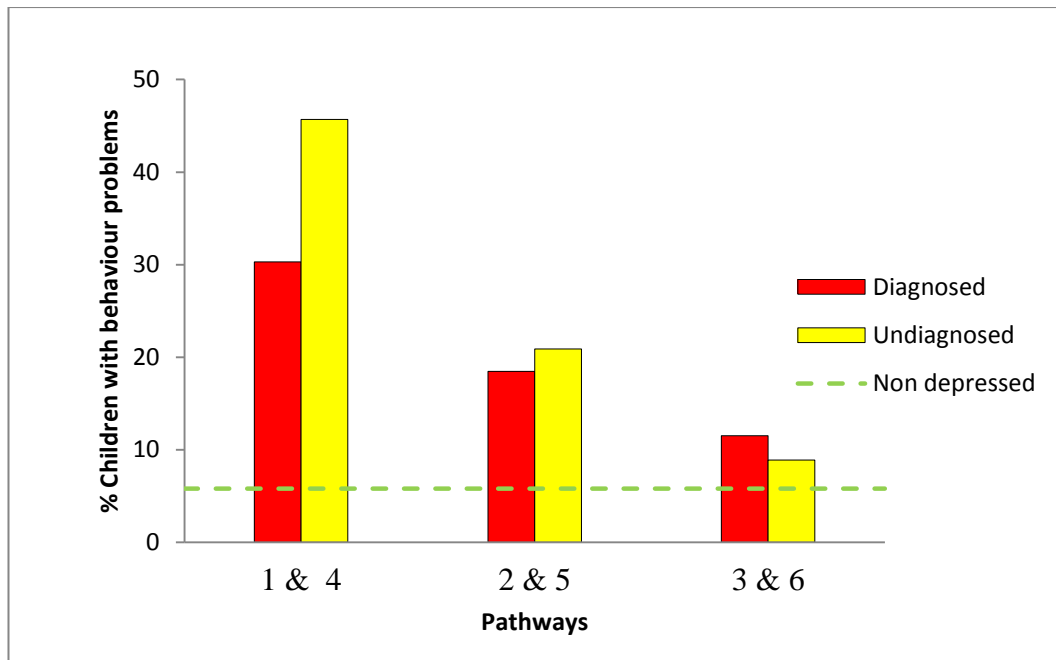
Figure 5.5. Proportions of children scoring above threshold for behavioural problems on the SDQ at age 7 according to the 7 pathway model

Although the raw proportions suggest that there may be differences between groups for child outcomes these could be due to the effects of confounding variables. As with model 1, univariate and multivariate regressions were carried out comparing each of the individual pathways to the non-depressed group and the comparable pathways to each other. These were conducted on both the continuous score and the dichotomised score as in Chapter 4.

Table 5.10. The effect of 6 maternal depression pathways on child total SDQ scores at age 7: Comparison to children of non-depressed mothers

	Univariate			Multivariate		
Pathway	Coef	95% CI	P	Coef	95% CI	P
D^D Continuous	0.83	0.72-0.94	<.0001	0.64	0.55-0.74	<.0001
D^D Intermittent	0.48	0.41-0.55	<.0001	0.40	0.34-0.47	<.0001
D^D Recovered	0.17	0.09-0.24	<.0001	0.10	0.03-0.18	0.003
D^U Continuous	1.09	0.96-1.24	<.0001	0.81	0.69-0.93	<.0001
D^U Intermittent	0.57	0.5-0.64	<.0001	0.44	0.38-0.51	<.0001
D^U Recovered	0.23	0.16-0.31	<.0001	0.19	0.13-0.26	<.0001

Code: D^D – diagnosed depression, D^U – undiagnosed depression

Table 5.11. The effect of 6 maternal depression pathways on children's odds of behavioural problems at age 7: Comparison to children of non-depressed mothers

| | Univariate | | | Multivariate | | |
Pathway	OR	95% CI	P	OR	95% CI	P
D^D Continuous	8.08	6.66-9.81	<.0001	6.39	5.22-7.82	<.0001
D^D Intermittent	3.57	2.97-4.31	<.0001	3.12	2.55-3.82	<.0001
D^D Recovered	2.23	1.70-2.93	<.0001	1.94	1.45-2.6	<.0001
D^U Continuous	11.28	9.00-14.15	<.0001	7.67	5.99-9.81	<.0001
D^U Intermittent	4.27	3.57-5.11	<.0001	3.28	2.69-4.00	<.0001
D^U Recovered	1.85	1.39-2.44	<.0001	1.60	1.20-2.12	0.001

Code: D^D – diagnosed depression, D^U – undiagnosed depression

As expected based on the results in model 2, children of mothers following any of the depression pathways were significantly more likely to score higher scores on the SDQ and be categorised as having behavioural problems. The results from the logistic regression mirrored the associations shown in the raw data where the odds of having a child that scores above threshold for behavioural problems increases with the type of depression pathway, with children of only postnatally depressed women having the lowest likelihood of scoring above threshold, children of intermittently depressed mothers being more likely and children of continually depressed mothers being the most likely.

Odds ratios differed little between the comparable pathways, as would be expected from the raw data, and from results presented in model 2.

Once groups of women were reduced into a single 'intermittent' category children were around 3 times more likely than children of non-depressed mothers to score above threshold for depression irrespective of whether depression was diagnosed or not.

Table 5.12. The effect of 6 maternal depression pathways on child total SDQ scores at age 7: Comparing children of undiagnosed mothers to children of diagnosed mothers

| | Univariate | | | Multivariate | | |
Pathway	Coef	95% CI	P	Coef	95% CI	P
D-Cont	0.15	0.05-0.25	0.003	0.10	0.01-0.20	0.029
D-Int	0.06	0.00-0.13	0.054	0.02	0.03-0.09	0.449
D-Rec	0.06	-0.03-0.15	0.193	0.08	0.08-0.17	0.079

Code: D-depressed (undiagnosed compared to diagnosed)

Although 3 of the comparable pathways showed significantly different results in model 1, after reducing the number of pathways to 1 'intermittent' group, no significant differences were found in child behaviour outcomes between the diagnosed and undiagnosed groups in the multivariate regression. This finding was further replicated in the logistic regression analysing the dichotomised scores (Table 5.13).

Table 5.13. The effect of 6 maternal depression pathways on children's odds of behavioural problems at age 7: Comparing children of undiagnosed mothers to children of diagnosed mothers

	Univariate			Multivariate		
Pathway	OR	95% CI	P	OR	95% CI	P
D-Cont	1.40	1.05-1.85	0.020	1.20	0.89-1.61	0.229
D-Int	1.20	0.94-1.52	0.143	1.05	0.81-1.37	0.705
D-Rec	0.83	0.57-1.21	0.327	0.82	0.56-1.22	0.328

Code: D⁻depressed (undiagnosed compared to diagnosed)

5.3.4 Model 4

Model 4 sought to ensure the validity of the results of model 3. When categorical variables are condensed, the clarity and richness of the data can be reduced. As an incremental effect was found in model 3 whereby the odds of children presenting with behavioural problems on the SDQ scale at age 7 increased in line with the persistency of the mothers depression, it is hypothesised that there may be a cumulative effect of exposure to episodes of depression.

In Model 3 an 'intermittent' category was derived to compress pathways. One problem with this is that mothers who reported depression at 2 time points are grouped with mothers who reported depression at 3 time points. According to the principles of the life course approach discussed in Chapter 2, the more frequent the exposure to negative stimuli the greater the effects may be, thus it is important to consider the number of exposure time points. Model 4 thus sought to expand the 'intermittent' category back out by categorising mothers according to the number of exposures reported. These could be in any order after the first sweep during the post natal period (9 months).

Women were thus separated into those identified as having 1, 2, 3, or 4 episodes of depression, 1 of which must be during the first postnatal year at sweep 1. This resulted in 9 pathways (Figure 5.6) described in Table 5.14.
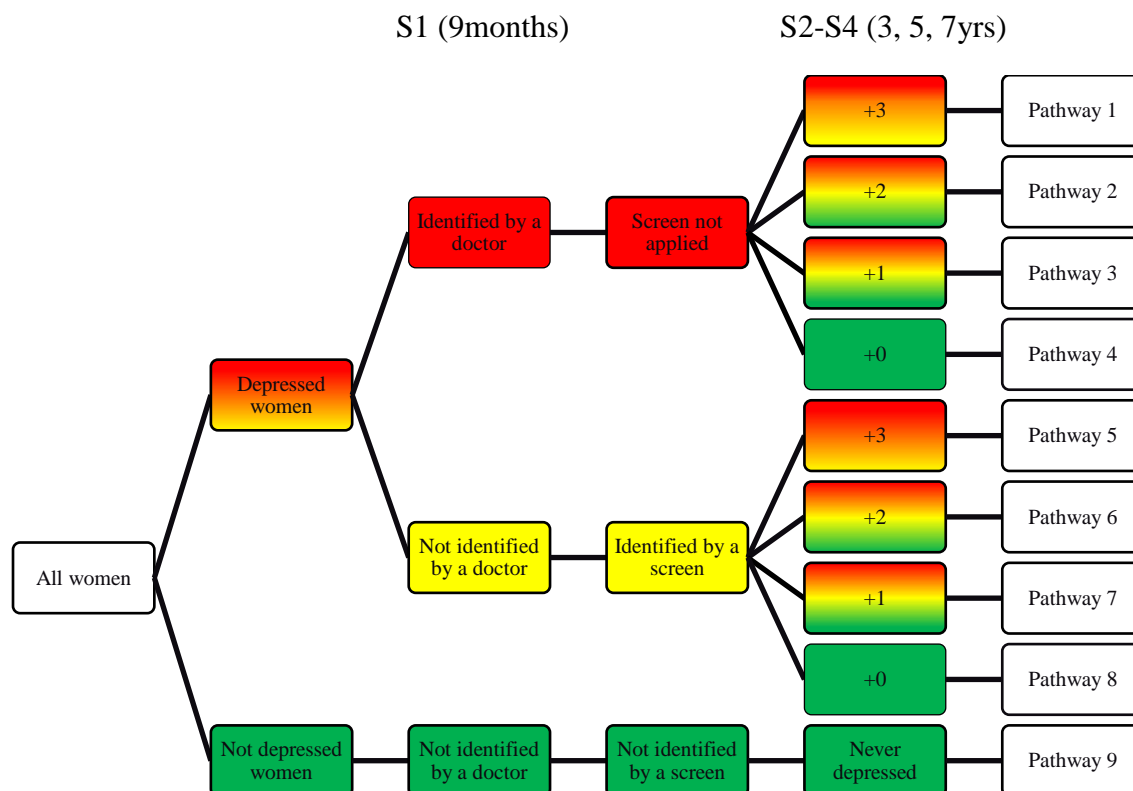
Figure 5.6. Decision model of depression over time mapping 9 pathways

Table 5.14. Proportion of mothers following each of the 9 pathways and associated proportions of child behaviour problems at age 7

Pathway	Code	N mother-infant pairs	N Children with behaviour problems	% Children with behaviour problems
1	D^D plus 3	198	60	30.30
2	D^D plus 2	143	31	21.68
3	D^D plus 1	144	22	15.28
4	D^D plus 0	183	21	11.50
5	D^U plus 3	116	53	45.70
6	D^U plus 2	120	36	30.00
7	D^U plus 1	172	25	14.53
8	D^U plus 0	214	19	8.90
9	Not depressed	4950	287	5.80

Code: D^D – diagnosed depression, D^U – undiagnosed depression

Table 5.14 demonstrates the incremental effect of multiple episodes of depression. In the diagnosed group more women remain categorised as depressed at all 4 sweeps than fully recover or are classified as depressed at fewer time points. Undiagnosed mothers appear more likely to spontaneously recover with the greatest proportion of undiagnosed women at sweep 1 being classified as not depressed at all 3 further time points.

As shown in Table 5.15, unadjusted results suggest that as the number of episodes of maternal depression increase the probability of children showing signs of behavioural problems increases. The effects appear to be most notable when the mother reports undiagnosed depression at sweep 1 in addition to depression at 2 or 3 other time points up to age 7.

Table 5.15. Comparable pathways of depression according to diagnosis from the 9 pathways model

	Diagnosed		Undiagnosed		
Pathway	% Mothers	% Children with behaviour problems	% Mothers	% Children with behaviour problems	Difference in behaviour problems (U-D)
D Plus 3	29.64	30.30	18.65	45.69	15.39
D Plus 2	21.41	20.26	19.29	30.00	9.74
D Plus 1	21.56	15.28	27.65	14.53	-0.75
D Plus 0	27.40	11.48	34.41	8.88	-2.6

Code: D- depressed

The results presented in Table 5.15 show that when depression occurs at only 1 time point, children of diagnosed women may be slightly more likely to develop behavioural problems compared to children of undiagnosed depressed women. Where depression is more frequent or consistent children of undiagnosed depressed women are more likely to develop behavioural problems. This reaffirms the notion that there is a relationship between frequency/duration of depression and child behaviour and also suggests that despite diagnosed women having an equal number of sweeps where depression was identified, receiving early treatment may be beneficial for children.

As can be seen in Figure 5.7 the main differences between children of diagnosed and undiagnosed depressed women can be observed when depression is reported at more time

points. Even when mothers are only depressed at sweep 1 children are still more likely to display behavioural problems than controls.
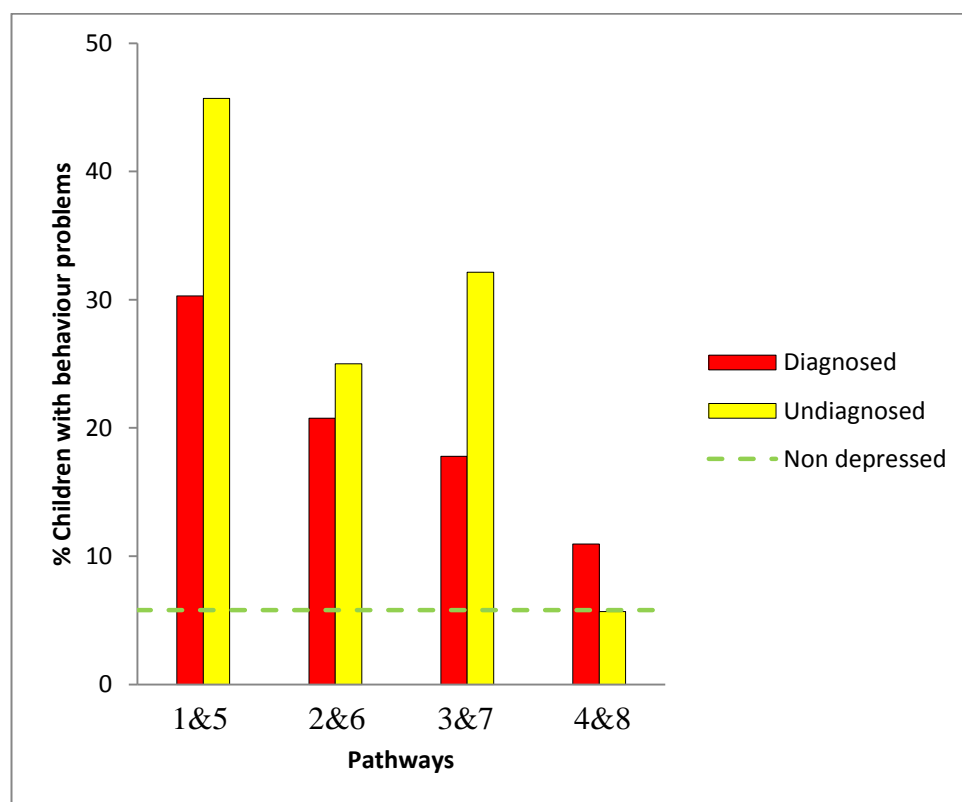
Figure 5.7. Proportions of children scoring above threshold for behavioural problems on the SDQ at age 7 according to the 9 pathway model

Again the results reported above are only based on raw associations between variables. There hypotheses were again tested using regression analysis to identify whether these differences were significant after taking into account the potential effects of confounding variables.

Table 5.16. The effect of 9 maternal depression pathways on child total SDQ scores at age 7: Comparison to children of non-depressed mothers

Pathway	Univariate			Multivariate		
	Coef	95% CI	P	Coef	95% CI	P
D^D plus 3	0.83	0.72-0.94	<.0001	0.65	0.55-0.75	<.0001
D^D plus 2	0.59	0.5-0.69	<.0001	0.51	0.42-0.60	<.0001
D^D plus 1	0.36	0.27-0.46	<.0001	0.30	0.21-0.40	<.0001
D^D plus 0	0.17	0.09-0.24	<.0001	0.10	0.03-0.18	0.004
D^U plus 3	1.09	0.96-1.24	<.0001	0.80	0.69-0.92	<.0001
D^U plus 2	0.84	0.73-0.95	<.0001	0.62	0.53-0.72	<.0001
D^U plus 1	0.42	0.34-0.51	<.0001	0.34	0.26-0.42	<.0001
D^U plus 0	0.23	0.16-0.31	<.0001	0.20	0.13-0.26	<.0001

Code: D^D – diagnosed depression, D^U – undiagnosed depression

Table 5.17. The effect of 9 maternal depression pathways on children's odds of behavioural problems at age 7: Comparison to children of non-depressed mothers

Pathway	Univariate			Multivariate		
	OR	95% CI	P	OR	95% CI	P
D^D plus 3	8.08	6.66-9.81	<.0001	6.40	5.23-7.83	<.0001
D^D plus 2	4.21	3.3-5.37	<.0001	3.59	2.77-4.66	<.0001
D^D plus 1	2.94	2.25-3.82	<.0001	2.64	1.96-3.54	<.0001
D^D plus 0	2.23	1.7-2.93	<.0001	1.94	1.45-2.60	<.0001
D^U plus 3	11.28	9-14.15	<.0001	7.68	6.00-9.82	<.0001
D^U plus 2	6.62	5.21-8.41	<.0001	4.45	3.38-5.86	<.0001
D^U plus 1	3.03	2.34-3.91	<.0001	2.59	1.98-3.40	<.0001
D^U plus 0	1.85	1.39-2.44	<.0001	1.60	1.20-2.13	0.001

Code: D^D – diagnosed depression, D^U – undiagnosed depression

Analysis on both the total SDQ scores and the dichotomised scores show a clear effect of accumulation. As the number of time points the mother is identified as been depressed increases both the child's SDQ total score and the likelihood of them scoring above threshold for behavioural problems increases. This is true for children of both diagnosed and undiagnosed mothers. The greatest effect sizes and odds ratios are seen for children of mothers with persistent depression that was not diagnosed during the first postnatal year.

As in previous models there seems to be an interaction between persistency of depression and diagnosis, where children of undiagnosed mothers who recover earlier fare better than comparative children of diagnosed mothers, but children of undiagnosed mothers with persistent depression fare worse than comparative children of diagnosed mothers.

Regression analyses were again performed on comparable groups to test the interaction between early diagnosis and persistency of depression and the effect this has on child behaviour.

Table 5.18. The effect of 9 maternal depression pathways on child total SDQ scores at age 7: Comparing children of undiagnosed mothers to children of diagnosed mothers

Pathway	Univariate			Multivariate		
	Coef	95% CI	P	Coef	95% CI	P
D Plus 3	0.15	0.05-0.25	0.003	0.10	0.01-0.20	0.028
D Plus 2	0.16	0.07-0.26	0.001	0.08	-0.01-0.17	0.085
D Plus 1	0.04	-0.05-0.14	0.385	0.03	-0.06-0.13	0.566
D Plus 0	0.06	-0.03-0.15	0.193	0.08	-0.01-0.18	0.078

Code: D-depressed (undiagnosed compared to diagnosed)

Table 5.19. The effect of 9 maternal depression pathways on children's odds of behavioural problems at age 7: Comparing children of undiagnosed mothers to children of diagnosed mothers

Pathway	Univariate			Multivariate		
	OR	95% CI	P	OR	95% CI	P
D Plus 3	1.40	1.05-1.85	0.020	1.20	0.89-1.61	0.228
D Plus 2	1.57	1.13-2.18	0.007	1.24	0.86-1.78	0.242
D Plus 1	1.03	0.72-1.47	0.865	0.98	0.67-1.45	0.933
D Plus 0	0.83	0.57-1.21	0.327	0.82	0.56-1.22	0.332

Code: D-depressed (undiagnosed compared to diagnosed)

As expected, very little difference was found in child outcomes between children in the two maternal depression diagnosis groups. Again only the difference in total SDQ scores between children of continuously depressed mothers was significant.

5.4 Discussion

The analysis in this chapter aimed to build on the results of Chapter 4, in which differences in child behavioural development were found between children of mothers with diagnosed or undiagnosed depression during the first postnatal year. These differences were shown to be significant at ages 3 and 5 but not at age 7.

The analysis in Chapter 4 assessed the hypothesis that depression, especially undiagnosed and untreated depression during the first postnatal year, will have a lasting impact on child development. This hypothesis is in line with critical periods theory, one of the models forming the life course approach, as described in Chapter 2. This model suggests that certain periods during the life course are more sensitive to negative circumstances than others and the first postnatal year has been identified as one such period in which exposure to maternal depression which can have lasting effects on children's development.

Whilst the findings from Chapter 4 do support a critical periods model in showing that children of mothers with either type of depression score are more likely to score above threshold for behavioural problems at ages 3, 5 and 7 compared to children of non-depressed mothers, intervention during this critical period through identifying and treating mothers depression is not sufficient to yield lasting positive effects shown by the equal

likelihood of children of undiagnosed and diagnosed mothers being equally likely to score above threshold for problems by age 7.

These findings prompted the hypothesis that changes in maternal depression over time may also be important.

This chapter thus tested accumulation theory also described in Chapter 2 which notes that although critical periods are important, multiple exposures to negative circumstances can have an additive effect on outcomes (Ben-Shlomo and Kuh, 2002, Najman et al. (2010).

Although there is some research that has explored the effects of persistent maternal depression on child outcomes (Brennan et al., 2000, Kiernan and Mensah, 2009) this has not differentiated between the effects of diagnosed and undiagnosed depression.

As no relationship was found between maternal depression and child cognitive outcomes in the analysis in Chapter 4, only the effects on behavioural development were explored.

5.4.1 Principal findings

A series of models were developed mapping changes in maternal mental health over time. This was an iterative process to derive a model most reflective of the pathways women actually take based on the data available.

For the purpose of addressing the analytical aims, 6240 mother-child pairs were included whereby mothers followed a depression pathway or reported no depression at any time point, and SDQ data at age 7 were available. Of these women 10.7% were reported diagnosed depression at sweep 1, 10.0% reported undiagnosed depression, and 79.3% reported no depression at any time.

When maternal depression was mapped into 17 pathways children of mothers following any of the depression pathways scored higher than children of non- depressed mothers and were more likely to score above threshold for behavioural problems. Effect sizes and odds ratios varied between pathways, with some having much greater effects than others.

Comparable pathways of depression were drawn where the only difference was whether or not the mothers depression had been diagnosed or not at sweep 1.

A sequential series of regression models was carried out taking an informed decomposition modelling approach. Regression analyses were conducted on both the continuous form of the SDQ and on the dichotomous form. The results showed children of undiagnosed mothers score equal to or slightly higher than children of diagnosed mothers regardless of future depression. Although some comparable pathways showed significant differences in SDQ scores between children of diagnosed and undiagnosed women, only 1 pathway significantly affected children's likelihood of scoring above threshold for behavioural problems. Children of mothers with undiagnosed depression at sweep 1 combined with depression at sweeps 2 and 4 were 4.1 times more likely to score above threshold for behaviour problems than children of mothers with comparable depression pathways but with diagnosis at sweep 1.

It could be argued that this is due to a cumulative effect however there was no difference between children when mothers were diagnosed at all 4 time points opposing this theory.

One possible explanation could be the small sample sizes of women following certain pathways. Within both undiagnosed and diagnosed groups of mothers the most common individual pathways followed are the 'recovery' pathways and the 'continuous pathways'. 34.4% of undiagnosed mothers fully recover with 18.6% reporting continuous depression. 27.4% of diagnosed mothers fully recover with 29.6% reporting continuous depression.

In an attempt to increase statistical power and thus reduce the chance of error a second model was constructed separating women into 'continuous' 'intermittent' or 'recovered'. This resulted in 7 pathways.
Regression analysis carried out on this model showed that when pathways are compressed to form an 'intermittent depression' group no differences are observed in child behavioural outcomes between children of undiagnosed or diagnosed mothers.

Although compressing the groups in this way leads to larger sample sizes, it does not allow for testing of the accumulation model as mothers with 2 or 3 reports of depression are merged together.

Consequently the third model was constructed testing the effects of cumulative depression. These groups were not formed on a time dependent basis, just on the number of sweeps mothers reported depression.

A clear effect of accumulation was found with children scoring higher and being more likely to score above threshold for behaviour problems on the SDQ for each additional sweep depression was reported. This effect was found for both undiagnosed and diagnosed mothers-infant pairs.

When mothers were depressed at sweep 1 children of undiagnosed mothers were 1.6 times more likely than children of non-depressed mothers to score above threshold for behavioural problems. This likelihood was increased to 1.9 for children of diagnosed mothers however logistic regression comparing the undiagnosed and diagnosed groups showed there was no significant difference by diagnosis.

This effect was reversed for persistent depression where children of undiagnosed mothers depressed at all 4 time points were 7.7 times more likely than children of non-depressed mothers to score above threshold for behavioural problems, compared to 6.4 times more likely for children of diagnosed mothers. Again no significant difference was found between the undiagnosed and diagnosed mother-infant pairs.

5.4.2 Comparison to previous literature

Previous work has shown an effect of persistent depression on child outcomes (Brennan et al., 2000, Ashman et al., 2008, Kiernan and Mensah, 2009). The present findings are in line with this showing a clear accumulation effect where an increase in exposure to maternal depression appears to be linked with an increase in SDQ scores and an increase in the likelihood of scoring above threshold for behavioural problems. The results of the present analysis go further to build on the findings of previous work by looking at differences in depression over a longer time period and by comparing changes according to diagnosis status.

The MCS has previously been used to investigate the effect of persistent maternal depression and child development. Kiernan and Mensah (2009) examined the effect of depression at sweeps 1 and 2 on child behaviour at sweep 2. In this case maternal depression was by self-report only based on answers to only 1 question enquiring about the mothers' feelings. When depression was only reported at sweep 1 children were 1.5 times more likely than controls to show behavioural problems, when depression was reported at both sweeps children were 3.74 times more likely than children of non-depressed mothers to show behavioural problems.

The present analysis showed similar findings for the 1 episode of depression only mother-child pairs, with children of depressed mothers being 1.6-1.9 times more likely than controls to show behavioural problems (depending on maternal diagnosis). The present analysis shows a stronger effect of persistent depression however this is likely to be due to the extended time course and the increased number of potential exposures.

5.4.3 Strengths and limitations

5.4.3.1 Strengths

This is the first analysis that has explored the longitudinal impacts of undiagnosed maternal depression during the postnatal year according to future changes in depression. Previous research on the effects on children of persistent maternal depression has been limited. The present analysis is able to show the pathways mothers depressed during the first postnatal year follow over time according to whether or not their depression was diagnosed and treated, whilst also showing the effects this has on the behaviour of their children.

A further strength of this analysis is its ability to empirically test both the critical periods and accumulation models drawn from life course theory, lending support for both models.

This analysis uses a large scale nationally representative data set. The millennium cohort study provides a large, contemporary, and highly representative UK data set in which to explore the effect of maternal depression on child behavioural development over time.

Analysing both continuous and dichotomous SDQ scores adds strength to the quality of analysis as both measures have been indicated as good representations of child behaviour (Goodman, 2001, Goodman and Goodman, 2009).

5.4.3.2 Limitations

Despite the strengths of the analysis, using secondary data does impose some limitations. Firstly, as mention in Chapter 4, the measures of maternal depression are not ideal though can be used to provide a good estimate of mental health and have been used in comparable work (Bartley et al., 2004, Mensah and Kiernan, 2010, Kiernan and Huerta, 2008).

Another problem encountered during this analysis is that it was not feasible to examine all pathways according to whether mothers were receiving treatment for depression at each time point due to small sample sizes.

As noted in Chapter 4, the MCS does not supply information on the type of treatment being received so it is not possible to examine the differences in maternal depression over time according to treatment type to assess whether certain treatments are more effective at resolving symptoms over the long-term and how this affects children.

5.4.4 Implications and further research

5.4.4.1 What does this study add?

Compared with previous studies the present analysis examines the effect of maternal depression over a longer time period (up to 7 years) and according to whether depression was missed by doctors and thus untreated or was diagnosed and treated during the first postnatal year.

The present analysis also lends support for both the critical periods and accumulation model showing that depression during the first postnatal year alone can have lasting effects on behaviour up to age 7 and this can be further accentuated by further depression in the future.

Through this analysis important information on mother's future mental health after childbirth has also been revealed, showing that maternal depression often persists over time regardless of whether it was treated suggesting that treatments as a whole received by MCS women are not collectively successful at completely curing depression. Further work is required to differentiate between the effectiveness of different treatments. It may be that no treatment can reliably prevent recurrence of depression and thus health professionals could consider checking up on mothers mental health beyond the first postnatal year.

5.4.4.2 Implications

These results demonstrate the differential effects of different pathways of maternal depression over time on child outcomes. A clear effect of accumulation is shown where more persistent depression is associated with greater effects on child behaviour. Very little benefits are observed for treating maternal depression in terms of child behavioural outcomes, with children whose mothers received treatment scoring only marginally better than children whose mothers were not treated. A large proportion of mothers who received treatment also reported depression at a further time point (72.6%) suggesting that current treatment practices may not be effective and completely treating or preventing future episodes of depression.

5.4.4.3 Further research

This analysis has highlighted several areas where further research is required. Firstly, it is clear that treating women for depression is not providing lasting benefits for children, who are performing significantly worse than children of depressed mothers who did not receive treatment and children of non-depressed mothers on behavioural measures. Research needs to be carried out to identify a suitable treatment that is successful in both treating maternal symptoms and improving the mother-child interaction which is likely to be responsible for the increase in child behaviour problems.

5.4.5 Data for the decision model

The primary objective of this thesis is to readdress the question of whether screening for maternal postnatal depression could be cost-effective given the effects it has on child outcomes, particularly those associated with undiagnosed depression. In Chapter 7 a decision model is devised to address this question by building upon a previous model published in a recent HTA report (Hewitt et al., 2009) as discussed in Chapter 1. The values used to populate this model will be calculated using the results 17 pathway model which allows for the most in depth exploration of the association between maternal postnatal depression diagnosis and long-term child outcomes, accounting for changes in maternal mood over time and will allow for closest adaptation of the original HTA model.

5.5 Conclusions

Although depression during the first postnatal year alone is sufficient to cause lasting effects on child development, supporting the critical periods model, further episodes of maternal depression over the early years accentuate this effect further. As the persistency of depression increases the negative impact maternal depression has on children's behavioural development also increases. Diagnosing and treating women for depression during the postnatal year has only small effects on children, who appear to perform equally poorly as children of mothers whose depression was not diagnosed and treated. These findings suggest the need for better identification of depressed women and treatments that are effective at not only treating maternal symptoms during the first postnatal year, but are effective in preventing future episodes of depression and are beneficial for child outcomes. Treatment options could thus potentially take a 'whole family' approach, targeting not only maternal symptoms but also family interactions. More research is required to investigate what treatment methods have the best outcomes for both the mother and the child.

Chapter 6: Exploring the long-term consequences of early behaviour problems: A Scoping review

6.1 Introduction

Earlier in this thesis postnatal depression was demonstrated to have long-term effects on children's behaviour until age seven. This chapter aims to build on this finding by exploring the literature surrounding the long-term outcomes associated with early behavioural problems by means of a scoping review.

Postnatal depression has been shown in this thesis and in other studies to have detrimental effects on child development including cognitive development and moreover, behavioural and emotional development. Results documented earlier in this thesis suggest that around 18% of new mothers will experience some level of postnatal depression. Children of these mothers, whether diagnosed or not, are more than twice as likely to exhibit signs of behavioural problems at age 7. Other studies have also noted the strong links between postnatal depression and behavioural problems in childhood.

Although the link between postnatal depression and the short term consequences on child behaviour are reasonably well established, less is known about how impaired behaviour during childhood affects children over the longer term.

Longitudinal outcomes can carry substantial weight in decisions of cost-effectiveness, particularly when taking a cost-benefit approach. James Heckman, a leading researcher in child development, is a consistent advocate of early intervention in the view of gaining later benefits and future cost-savings (Heckman, 2006, 2007, 2008). Such thinking has motivated the development of models translating outcomes in childhood to longer term outcomes in adulthood through extrapolation work. The most notable example of this is the Washington State Institute for Public Policy (WSIPP) in the US which was designed to predict the potential long-term impact of competing investment options for child wellbeing, as well as the costs and economic returns of interventions (Lee et al., 2012). The WSIPP cost-benefit model takes a 3-step approach; firstly, existing evidence is used to determine the effectiveness of programmes and approaches, secondly, costs and benefits are calculated producing a ranking of public policy options and finally, the risk associated

with the conclusions of the model output is measures by testing how outcomes may vary when estimates and assumptions change (Lee et al., 2012). In the UK, the Social Research Unit is working to adapt the WSIPP model into a UK context with the aim of providing policy makers with 4 key summary statistics, namely; a net present value, benefit-cost ratio, internal rate of return on investment, and a measure of risk associated with bottom-line estimates. Future outcomes are reported in terms of criminal activity, labour market earnings and educational attainment (Little et al., 2013). If a difference in child outcomes was found at age 7 between children of undiagnosed and diagnosed mothers, a similar model would be derived to extrapolate the long-term effects of early behaviour problems into adulthood outcome measures, similar to those highlighted in the WSIPP and SRU models. This review thus aims to identify the long-term outcomes associated with early behavioural problems as to identify potential avenues for future benefits should identifying mothers and providing appropriate treatment as a consequence of screening prove beneficial in reducing the likelihood of children developing behavioural problems.

6.1.1 Building trajectories: Early childhood behaviour and long-term outcomes

Researchers supporting the life course approach, as described in Chapter 2, theorise that circumstances in the early years set individuals on lifelong trajectories (Bartley, 2004, Braveman and Barclay, 2009). It is thought that early behavioural problems may have lasting consequences that impact over the life course affecting various important life outcomes such as academic attainment and the likelihood that individuals engage in criminal activity in adolescence and adulthood (Klein et al., 2012) which can in turn affect employment and ultimately labour market earnings (Knapp et al., 2011) potentially placing a financial burden on society.

Although behaviour problems can be age related and may spontaneously resolve in line with developmental stage, children who exhibit early signs of behaviour problems are likely to display behaviour problems over time (Hobcraft and Kiernan, 2010). Behaviour problems that resolve while the child is still in the early stages of development are unlikely to have lasting effects. Problems arise when behaviour issues persist beyond school entry. A recent article using data from the MCS noted that one in six primary school aged

children (16.8%) is now streamed by age seven (Hallam, 2012). In addition to this almost 26% of children were set for both literacy and maths and 11.2% were set for only maths (8.2%) or literacy (3%).

Several studies have noted that in some children behavioural problems in the early years may persist for a much greater time period and thus are more likely to affect other aspects of their life. Campbell and colleagues conducted a small scale study assessing the stability of early behavioural problems over a 10 year period following 46 children who had been identified by parents as inattentive, hyperactive and disruptive, and 22 controls. Following an initial assessment at age 3, further assessments were conducted at ages 4, 6, 9 and 13 years. By age 6, 50% of the 'problem' group continued to display signs of behavioural issues (Campbell et al., 1986). At age 9, 48% of the initial sample was clinically diagnosed as having a behavioural problem compared to only 16% of controls (Campbell and Ewing, 1990). Children who were shown to have continuous behaviour problems up to age 9 were shown to continue to do so through to age 13 meaning that almost half the children identified as behaviourally impaired at age 3 continued to be identified as such a decade later (Campbell, 1995). The small sample size used in this study demands that caution is applied to the interpretation of findings and larger sample sizes are needed to determine the meaningfulness of these results. Nonetheless Campbell and colleagues' work does point towards a lasting stability of behavioural problems from pre-school through to adolescence in a large portion of those children identified with early problems. Such findings have been replicated in other small scale studies over a similar time frame (Egeland et al., 1990, McGee et al., 1991, Fischer et al., 1990). This highlights an important issue in the persistency of early behavioural problems which is likely to have serious consequences for general development from childhood to adulthood.

Given that it was found earlier in this thesis that postnatal depression largely affects child behaviour, the purpose of this chapter is to explore the link between early behaviour problems and children's long term outcomes, primarily in terms of educational attainment, employment and adolescent/adult criminality as these aspects appear most prominent in the research as important outcomes. This information could be used to inform the decision model to extrapolate the effects of postnatal depression on children over a much longer time period into adulthood.

6.1.2 Aims

This chapter aims to examine the extent and findings of research activity relating to the long term effects of early behavioural problems on children's educational attainment, adolescent/adulthood criminal activity, and labour market activity as well as any other long term consequences identified. If a positive effect of identifying and treating maternal depression were to be identified, these findings could be used to provide estimates of the long-term effects of early child behaviour problems to inform the decision model presented in Chapter 7. This chapter thus aims to identify the potential avenues for future benefits in early intervention to prevent child behavioural problems. Findings could not only be used to inform the current question but also be used to extrapolate long-term outcomes associated with changes in early child behaviour problems for other research questions.

6.1.3 Questions

- What is the relationship between early behavioural problems and educational attainment?
- How are early behavioural problems related to criminal activity in later life?
- How are early behavioural problems related to labour market activity and earnings?

6.2 Method

6.2.1 Design

An interpretative scoping review was carried out to explore the study questions. This type of review is an alternative to a full systematic review and has been described as aiming '*to map rapidly the key concepts underpinning a research area and the main sources and types of evidence available*' (Mays et al., 2001). A systematic review typically focuses on a well-defined question where appropriate study designs such as RCT's can be identified in advance. In contrast to this, a scoping review tends to address broader subject areas where studies of different designs might be appropriate for inclusion. In addition to this, systematic reviews tend to aim to provide answers to questions from a relatively narrow range of quality assessed studies, which may not be the case in scoping studies which do

not generally adopt quality control measures for included studies. Arksey and O'Malley (2005) identify four circumstances in which a scoping review is appropriate;

1. To examine the extent, range and nature of research activity
2. To determine the value of undertaking a full systematic review
3. To summarise and disseminate research findings
4. To identify research gaps in the existing literature

The present study aims to examine the range and nature of research activity concerning the study questions and summarise findings using the framework outlined by Arksey and O'Malley (2005). According to Arksey and O'Malley the method adopted for identifying literature in a scoping study needs to achieve in-depth and broad results. This means that identified studies are not stratified by particular study designs as might be the case in systematic reviews. The scoping methodology aim is to identify all relevant literature regardless of study design; consequently strict limitations are not placed on search terms. The process requires an iterative and flexible approach using different combinations of key words rather than a definitive stringent search strategy.

6.2.2 Search strategy

Online data bases including CINAHL, Pubmed, Ovid and PsychInfo were searched. Searching techniques used included search tools such as medical subject headings (MeSH) and Boolean operators to expand and narrow searches. In the online databases single and combined search terms were used such as: child, behaviour, education, and crime. These searches were further refined using additional layers of terms which were more specific such as attainment, test scores, GCSE, A Level, criminal activity, antisocial behaviour. In addition to this various combinations of key words were entered into Google Scholar to uncover any additional publications or grey literature (see Appendix 4 for an example search strategy). As this was an iterative process rather than systematic, search terms were adapted to suit the literature. The initial search revealed few papers referring to 'behaviour' as an umbrella term. Literature was more likely to refer to the effect of particular behaviour such as conduct problem, and attention problems. The search was therefore modified to include the specific subgroups of behaviour rather than searching for articles described behaviour problems as a whole. Retrieved articles were assessed for relevance, with

irrelevant articles excluded. Bibliographies of relevant articles were hand searched to locate any further suitable material.

6.2.3 Quality Assessment of studies

A primary study's 'scientific quality' is highly important for determining the external and internal validity of its results. For studies that do not follow the methodological rigour of randomised controlled trials there is no one widely accepted evaluative tool for study quality. As a result of this, a Health Technology Assessment (HTA) Programme Report was carried out in 2003 by Deeks and Colleagues to identify and evaluate available quality assessment tools and checklists for non-randomised data (Deeks et al., 2003). A succession of systematic reviews and empirical investigations were carried out using a wealth of data from various sources. Results showed that 193 individual quality tools were identified, and although 14 were noted to be of high quality, only 6 were deemed suitable for use in a systematic review.

Consequently, the assessment tool to be used for the present work was selected from these 6. The 'Quality Assessment Tool for Quantitative Studies' (Thomas, 1998) was developed to assess the methodological quality of studies and includes 21 items evaluating: selection bias, study design, confounders, blinding, data collection methods, withdrawals and dropouts, intervention integrity and analysis. The tool then requires the reviewer to allocate an overall rating of the study (strong/moderate/weak) in order to provide a comparable quality rating. This tool was selected due to the scope of items covered and because it is able to deal with a range of study designs including both randomised and non-randomised studies (Deeks et al., 2003). In addition to this Thomas's method has been shown to have good construct validity and test-retest reliability and has performed favourably in the assessment of non-randomised data (Thomas et al., 2004), though less well for assessments of randomised data (Armijo-Olivo et al., 2012). As the studies relevant to this topic are unlikely to be RCT's this tool is deemed appropriate.

6.3 Findings

This section presents the findings of the scoping review.

6.3.1 Early childhood behaviour and educational attainment

Table 6.1. An overview of the included studies including quality assessment rating relating to educational outcomes

Author, year, title	Aim	Sample	Measures & age of assessment	Main findings	Quality rating
Breslau et al. (2009) The impact of early behaviour disturbance on academic achievement in high school.	To investigate the effect different types of behaviour disturbance have on academic attainment	693 ethnically diverse families from a cohort in the US	**Age 6** Behaviour: Achenbach Teacher Report Form (TRF) **Age 17** Educational attainment: Woodcock Johnson-Revised (WJ-R) measuring reading and maths skills	Significant association between behaviour and both maths and reading assessments (p<.001). Only the effect of attention problems was maintained when controlling for other factors. Coefficient: -0.1 p<.05 Reading -0.1 p<.05 Maths	Moderate
Duncan et al. (2007) School readiness and later achievement.	To estimate links between school entry attention and socio-emotional skills with later reading and maths achievement.	6 longitudinal data sets, 4 US based, 1 UK based and 1 Canada based.	**Various ages** Behaviour: various parent **Various ages (9-13)** Educational attainment: Various measures of reading and maths skills	Meta-analytic results indicate that attention skills at school entry predict subsequent reading and maths achievement. Attention skills coefficient: .08 p<.001 Reading .11 p<.001 Maths .10 p<.001 Combined No significant differences were found for other externalising behaviour or internalising behaviour types.	Moderate

Table 6.1. An overview of the included studies including quality assessment rating relating to educational outcomes ctd.

Klein et al. (2012) Clinical and functional outcome of childhood attention-deficit/hyperactivity disorder.	To examine whether children identified as ADHD at age 8 have worse educational, occupational and economic outcomes.	135 men diagnosed with ADHD at age 8 and 136 controls	**Age 8** Behaviour: clinical diagnosis **Age 41** Educational attainment: self-report of qualifications	On average men with ADHD had 2 ½ fewer years of schooling than controls. 31% did not complete high school (4.4%) controls. 3.7% had a higher degree (29.4% controls). Both these differences were significant, p<.001.	Weak
Mcleod and Kaiser (2004) Childhood emotional and behavioural problems and educational attainment.	To investigate whether childhood emotional and behavioural problems reduced the likelihood of graduating high school and attending college.	US sample of 424 families from the NLSY data set were used.	**Age 6-8** Behavioural problems: Child Behaviour Checklist (CBCL). **Age 20-22** Educational attainment: Self-reported high school completion and college enrolment	When participants were split into highest rates of behaviour problems (upper 10th percentile) versus controls, externalising problems accounted for a 31% reduction in the likelihood of obtaining a high school degree (p<.01), and internalising problems leading to a 26% reduction (p<.05). Only externalising problems were significant predictors of university enrolment, p<.05.	Strong
Woodward and Fergusson (2000) Childhood peer relationship problems and later risks of educational under achievement and unemployment.	To investigate associations between teacher-rated peer relationship problems at age 9 and subsequent risks of education under achievement and employment by age 18.	964 respondents from CHD study in New Zeland.	**Age 9** Behaviour: Teacher interviews **Age 18** Educational achievement: attainment of A-C grades in School Certificates	Children with behavioural problems attain 1 fewer pass than non-problemed (p<.01) Children with poor behaviour were also more likely to report conflict with teachers (12.4% vs 4%), be suspended from school (21.5% vs 1.8%) and to report leaving school at age 16 or younger (40.5% vs 14.6%).	Moderate

Table 6.1. An overview of the included studies including quality assessment rating relating to educational outcomes ctd.

Polderman et al. (2010) A systematic review of prospective studies on attention problems and academic achievement.	A systematic review of literature linking attention problems to academic outcomes	16 studies of varying sample sizes	**Age 4+** Behaviour: various parent reported measures **Age up to 25** Educational attainment: various	Early attention problems were significantly related to academic attainment however measured even after accounting for confounders.	Moderate
Frazier et al. (2007) ADHD and Achievement: Meta-Analysis of the Child, Adolescent, and Adult Literatures and a Concomitant Study With College Students	To carry out a meta-analysis of evidence linking ADHD to academic achievement.	72 studies were included with varying sample sizes.	**Age not reported** Behaviour: Clinical diagnosis of ADHD **Age not reported** Measures of academic achievement varied between studies.	Diagnosed ADHD during childhood was significantly related to academic attainment, with attention skills appearing most predictive.	Weak

Table 6.1. An overview of the included studies including quality assessment rating relating to educational outcomes ctd.

Massetti et al. (2007) Academic Achievement Over 8 Years Among Children Who Met Modified Criteria for Attention-deficit/Hyperactivity Disorder at 4–6 Years of Age	To assess academic attainment over an 8 year period in children who met criteria for ADHD	255 (130 controls) children	**Age 4-6** Behaviour: clinical diagnosis of ADHD **Age 12-14** Educational attainment: standardised measures of reading and maths skills.	ADHD was significantly related to academic achievement, particularly inattention. Internalising problems were also shown to be significantly related to academic achievement.	Moderate
Washbrook et al. (2013) Pre-school hyperactivity/attention problems and educational outcomes in adolescence: prospective longitudinal study	To investigate whether hyperactivity/ inattention and conduct problems are associated with academic outcomes at age 16.	11640 from the ALSPAC cohort study (UK)	**Age 3** Behaviour: SDQ **Age 16** Educational attainment: GCSE attainment	Hyperactivity problems: Boys failure to achieve 5 A*-C OR 1.57 Girls failure to achieve 5 A*-C OR 1.47 Conduct problems: Boys failure to achieve 5 A*-C OR 1.39 Girls failure to achieve 5 A*-C OR 1.46	Strong

An increasing body of literature has shown that postnatal depression can have substantial consequences for child behavioural development (as described in Chapter 1). Problematic behaviour has in turn been linked with concurrent problems in learning which may go on to affect long-term educational trajectories. Studies have shown associations between behaviour and educational development in both directions with behaviour problems such as poor attention being predictive of educational attainment (Rabiner et al., 2000) and early educational attainment in areas such as reading being predictive of developing later behaviour problems (Halonen et al., 2006).

In this scoping study, eight studies and two reviews were identified that explored the relationship between early behaviour problems and later educational attainment.

Within these studies attainment was measured in several different ways: domain specific, (i.e. measuring particular facets of education such as reading or mathematical ability, number of years spent in education, highest qualification obtained, high school graduation and college enrolment (US studies)), or specific attainment in national qualifications, namely GCSE's (UK), or school certificates (New Zealand).

Five papers, including the two reviews, examined the relationship between early behaviour and attainment through domain specific assessments (Duncan et al., 2007, Breslau et al., 2009, Polderman et al., 2010, Frazier et al., 2007, Massetti et al., 2007).

The collective findings from these studies were largely similar, with early behaviour problems of any type being consistently significantly linked with later reading and mathematical attainment.

In one of these studies, data from six longitudinal data sets (4 American, 1 British, and 1 Canadian) were used to estimate the effects of behavioural problems (both externalising and internalising) at school entry with later school reading and mathematical achievement (Duncan et al., 2007). The timing of the outcome measures varied across studies ranging from age 9 to age 14 years. All six data sets revealed strong links between early attention problems and later academic performance. Other behaviour types encompassed by the internalising and externalising types of behaviour were not significantly related to later

achievement. Patterns of association were similar for boys and girls and for children from high and low socio-economic backgrounds.

Similar findings were also noted in relation to children with diagnosed ADHD during early childhood (age 4-6 years) in a longitudinal 8 year follow up study (Massetti et al. 2007). Again early attention problems were found to be significantly related to future mathematical achievement. However, in contrast to the findings of Duncan and colleagues, hyperactivity (an externalising behaviour type) was also found to be significantly related to future reading attainment. This may be due to the diagnostic status of the children examined by Massetti et al. and potential increased severity of behavioural problems. It may be the case that educational attainment is only affected in the most problematic children.

Unlike the relatively short-term follow-ups carried out in these two studies, the final paper investigating the effects of early behaviour problems on specific domains of educational attainment. Breslau et al. (2009) follow children up to 17 years old from 6 years old. After adjusting for a range of possible confounders, early teacher ratings of attention, internalising and externalising problems at age 6 were shown to predict mathematical and reading achievement at age 17. Each of the behavioural problem areas was significantly associated with the mathematical and reading achievement scores, with one marginal exception where the P value for the coefficient relating externalizing behaviour with reading achievement was .06. This advances the previous two studies which linked attention and hyperactivity problems with later attainment, to include other behaviour types as well. Such findings highlight the importance of a range of behaviour types during early childhood for later educational achievement.

Of the two reviews, one carried out a meta-analysis of 72 studies (Frazier et al., 2007) whilst the other collated evidence using a 'best-evidence synthesis' approach of 16 studies due to large variation between independent and dependent measures (Polderman et al., 2010). This difference in methodology may reflect the different periods of time the included studies in each review encompass, as the review by Frazier et al., included studies published between 1990 and 2004, though the majority were issued pre-2000, whereas Polderman et al., targeted much more recent literature with the majority of included studies being published within the last 10 years and may be more representative of the increased

diversity of modern measures of academic attainment making standard meta-analytical processes difficult. Despite using different analytic strategies and included studies from largely different time frames, both reviews converged on similar results, finding in both cases that children with early behaviour problems are more likely than controls to perform poorly on standard tests of reading and mathematical ability. Both reviews noted the importance of confounding variables ensuring greater scientific confidence in the results; however Frazier et al. stratified results only according to children versus adolescents with actual age of participants within each included study being omitted.

In addition to examining the effects of behaviour on specific domains of education, Polderman et al. also extracted results from studies concerning the effects of behaviour on the likelihood of graduating high school and attending college (US studies only) though these were not commented upon, possibly because only two studies measuring attainment in this way were identified.

Both of these studies were said to have found a significant relationship between child behaviour problems and high school graduation and/or college enrolment (Barkley et al., 2006, Lambert et al., 1988) but fail to meet criteria for inclusion in the present review due to being outdated (Lambert et al., 1988) or failing to look explicitly at early childhood behaviour or stratify results by age (Barkley et al., 2006).

One study that met the requirements for the present review examines this relationship (McLeod and Kaiser, 2004). In this study behaviour was assessed at ages 6-8 years using the Behavioural Problems Index which identifies both externalising and internalising behaviours in a sample of 424 individuals from the US National Longitudinal Survey of Youth (NLSY) cohort. Results showed that both externalising and internalising problems were associated with the probability of receiving a high school degree. When participants were split into highest rates of behaviour problems (upper 10th percentile) versus controls, externalising problems accounted for a 31% reduction in the likelihood of obtaining a high school degree ($p<.01$), and internalising problems leading to a 26% reduction ($p<.05$). Neither externalising nor internalising behaviours were shown to be significantly predictive of University enrolment when all of the background variables were included in the model.

Using measurements of the number of years of schooling and the highest qualification level obtained, Klein et al. (2012) reported very similar results. Using data from a 33 year

follow up of men with childhood behaviour problems (teacher rated or psychiatrist diagnosed) identified at age 8, it was found that on average, men with childhood ADHD acquired 2 and a half fewer years of schooling than the comparison group. Of the men with childhood behaviour problems, 31.1% did not obtain a high school degree (age 18) compared to only 4.4% of comparison participants. In terms of overall qualification attainment, men with childhood behaviour problems were also much less likely to obtain a higher degree than control men, 3.7% compared to 29.4%.

Although these findings are largely in line with the studies noted above, Klein and colleagues completely omit women from their analysis making the findings less comparable with other studies. Nevertheless the results provide strong evidence that early behaviour problems can have an effect on long-term education.

Finally, two papers examined education in terms of standardised national test scores, one of which was carried out in the UK using grades on national GCSE's, the other in New Zealand using grades on School Certificates, which are equivalent to British GCSE's.

Both studies noted significant relationships between early behavioural problems and grade attainment at age 16, though each looked at slightly different aspects of behaviour.

In the New Zealand study Woodward and Fergusson (2000) examine the relationship between education attainment and peer relationship problems (a measure of internalising problems) at age 9 in a sample of 964 children as part of the Christchurch Health and Development Study.

Children's behaviour was measured through teacher interviews using a compilation of questions from two existing questionnaires (Rutter A Scale and Conners Teacher Rating Scale) as well as customised items. Univariate analyses showed striking differences were found between groups, with children who had no or low levels of problems attaining on average 4.09 passes (SD 1.96) compared to only 1.97 passes for the most problematic children (SD 2.2) $p < .0001$. However, these effects were reduced when confounders were controlled for. Multivariate analyses showed that the most disturbed 10% received on average 1 less certificate at grade A-C compared to the least disturbed 25%, though this difference remained significant at a high level ($p < .0001$).

In addition to problems with attainment the results also showed an association between the quality of children's early peer relationships and later interpersonal and school related difficulties. Compared to children with few or no peer problems, the most disturbed children (upper 10%) were more likely to report relationship problems with their teachers, 4% vs 12.3% (p<.0001), report being suspended from school between the ages of 13-16, 1.8% vs 21.5% (p<.0001), and report leaving school early (16 years or below) 14.6% vs 40.5% (p<.0001), showing that the long-term consequences of early behavioural problems extend beyond the scope of just grade attainment into school behaviour and relationship development.

In contrast to the work done in New Zealand which looked only at internalising behaviour types, Washbrook and colleagues' work in the UK examines the long-term effects of externalising behaviours measures around age 4 using the SDQ. Points were allocated to children's 8 best grades to create a continuous variable. 16 points were allocated for a G grade with 6 additional points being added with each grade increase up to a maximum of 58 points for an A* grade. A binary variable was also derived which categorised whether children had attained 5 A*-C grades (to include English and Maths). A wide range of variables were controlled for included family demographics, gender, and parental mental health during the early years.

SDQ scores for hyperactivity were classified as normal (0-5), borderline abnormal (6) or abnormal (7-10) in line with recommended cut-offs (Goodman, 1997). Boys but not girls classified as borderline abnormal on the hyperactivity subscale scored significantly less GCSE points than controls, with boys scoring 13.31 points less (p<.01) and girls scoring 6.79 points less (not significant). When levels of behaviour problems increased into the upper threshold classified as abnormal, both boys and girls scored significantly less GCSE points than controls, with boys scoring 20.9 points less and girls scoring 14.42 points less (p<.01).

SDQ scores for conduct were classified as normal (0-2), borderline abnormal (3) or abnormal (4-10) again in line with recommended cut-offs. Both borderline and abnormal levels of conduct disorder significantly affected GCSE scores for both girls and boys, with borderline boys scoring 12.08 less points than controls compared to 12.58 less points for

girls, and boys rated as abnormal scoring 16.94 less points than controls compared to 17.97 points less for girls (all comparisons, p<.01).

These differences were mirrored in the findings relating to the likelihood of obtaining 5 or more GCSE passes, though findings were only significant for the most abnormal children. For both conduct and hyperactivity problems both boys are girls were around 1 and a half times more likely to fail to obtain 5 good GCSE passes compared to controls (all comparisons <.01).

This is the first study of its kind to map early behaviour problems over such a large time scale to GCSE attainment at age 16.

As the majority of studies focus on only specific domains of education and are measured using varying tools at varying ages, the findings of these studies cannot be synthesised and integrated into the decision model concerning the question of whether screening and treating postnatal depression in the view to prevent behavioural problems from developing. Only the studies examining attainment on standardised national tests (Washbrook et al., 2013, Woodward and Fergusson, 2000) provide relevant and usable data that can be related to future costs and benefits.

6.3.1.1 Assumptions for the decision model: Educational attainment

Given the large differences in the methods used to educational attainment meta-analysis of study findings is not possible. Only 2 studies measured outcomes in terms of qualification attainment (Woodward and Fergusson, 2000, Washbrook et al., 2013), however only one of these was based in a UK context (Washbrook et al., 2013). Because of this and the strong methodology employed by the study, these results will be used as the effect sizes for long-term educational attainment in the decision model to estimate cost-effectiveness.

6.3.2 Early behaviour and later criminal activity

Table 6.2. An overview of the included studies including quality assessment rating relating to criminal activity

Author, year, title	Aim	Sample	Measures	Main findings	Quality
Stevenson and Goodman (2001) Association between behaviour at age 3 years and adult criminality.	To determine whether preschool behaviour problems increase the risk of later criminal convictions.	Diverse sample of 828 participants	**Age 3** Behaviour: parent interviews **Age 23-24** Criminal activity: Criminal records (Criminal Records Office)	Hyperactivity and conduct problems during childhood were shown to increase the chances of offending during adulthood after covariates applied: Hyperactivity: OR 1.86 p<.01 Conduct: OR 1.91 p<.01	Moderate
Babinski et al. (1999) Childhood conduct problems, hyperactivity-impulsivity and inattention as predictors of adult criminal activity.	To determine the contribution of early conduct problems as well as hyperactivity and inattention on offending in later life.	305 participants from San Francisco, US	**Age 9** Behaviour: Parents and teachers Children's Attention and Adjustment Survey (CAAS). **Age 26** Criminal activity: self-report and official records from California Department of Justice.	Hyperactivity and conduct problems during childhood were shown to increase the chances of offending during adulthood. Hyperactivity: OR 2.85 p = .015 Conduct: OR 7 p =.001. No controls were included in the model.	Weak

Table 6.2. An overview of the included studies including quality assessment rating relating to criminal activity ctd.

Simonoff et al. (2004) Predictors of antisocial personality: continuities from childhood into adult life.	To explore the independent and joint effects of childhood characteristics on the persistence of antisocial behaviour into adult life.	Sample of 225 men in retrospective study	**Age 7 or below** Behaviour: clinical diagnosis of hyperactivity and/or conduct problems **Age 38.2 (mean)** Criminal activity: official data from the Criminal Records Office.	Juvenile offences: Hyperactivity: OR 3.13 p<.001 Conduct: OR 11.52 p<.001 Adult offences (age 17+): Hyperactivity: OR 2.32 p<.05 Conduct: OR 3.36 p<.01	Moderate
Bor et al. (2004) Early risk factors for adolescent antisocial behaviour: an Australian longitudinal study.	To identify early risk factors for adolescent antisocial behaviour.	Sample of 5278 participants from Australian Longitudinal study	**Age 5** Behaviour: CBCL for parents **14 years** Criminal activity: maternal reports using the delinquency subscale of the CBCL	Antisocial behaviour: Aggression problems: OR 3.1 p<.01 Attention problems: OR 2.51 p<.01	Weak
Satterfield and Schell (1997) A prospective study of hyperactive boys with conduct problems and normal boys: Adolescent and adult criminality.	To examine the relationship between attention and hyperactivity problems in childhood and criminality in adolescence and adulthood.	89 men diagnosed hyperactive 87 control males.	**Age 6-12** Behaviour: clinical diagnosis of hyperactivity **Age 19-25** Criminal activity: official criminal records data were obtained	Juvenile criminal activity of hyperactive individuals compared to controls: Arrest rates: 46% vs 11% p<.001 Incarceration rates: 22% vs 1% p<.001 Adult criminal activity: Arrest rates: 21% vs 1% p<.001 Incarceration rates: 12% vs 0% p<.01. Juvenile crime significantly predicted adult crime	Moderate

Table 6.2. An overview of the included studies including quality assessment rating relating to criminal activity ctd.

Klein et al. (2012) Clinical and functional outcome of childhood attention-deficit/hyperactivity disorder.	To examine whether children identified as ADHD at age 8 have worse educational, occupational and economic outcomes.	135 men diagnosed with ADHD at age 8 as well as 136 controls	**Age 8** Behaviour: clinical diagnosis of hyperactivity and attention problems **Age 41** Criminal activity: self-report	Percentage of the sample reporting incarcerations of 1 day or more. ADHD: 36.3% Control: 11.8% $X^2 = 22.36$ p<.001	Weak
Murray et al. (2010) Very early predictors of conduct problems and crime: results from a national cohort study	To identify early predictors of later conduct disorder and crime.	16401 participants from the BCS70.	**Age 5** Behaviour: Rutter A2 behaviour questionnaire for parents **Age 34** Criminal activity: convictions were self-reported	Criminal convictions: Conduct: Boys: OR 1.6 p<.05 Girls: OR 2.0 p<.05 Internalising problems negatively predicts for boys only: Boys: OR 0.6 p<.05	Strong
Fergusson et al. (2005) Show me the child at seven: the consequences of conduct problems in childhood for psychosocial functioning in adulthood.	To examine the long term consequences of childhood conduct problems.	1265 participants from New Zealand cohort study	**Age 7-9** Behaviour: parents and teachers: Rutter A2 behaviour questionnaire and Connors' parent and teacher questionnaires **Age 25** Criminal activity:	Childhood behaviour was significantly related to crime. There was evidence of a gradient effect in terms of severity of behaviour symptoms. Rates of offending: 1-50th percentile: 4.5 51-80th percentile: 7.3 81-95th percentile: 12.3 96-100th percentile: 19.5 P<.0001	Moderate

Eight studies were identified that explored the relationship between early behaviour problems and later criminal activity.

Within these studies the measurement of criminal activity was relatively consistent. In six studies criminal activity was measured by convictions (Stevenson and Goodman, 2001, Babinski et al., 1999, Simonoff et al., 2004, Satterfield and Schell, 1997, Murray et al., 2010, Fergusson et al., 2005), these were either self-reported (n=3) or were sourced from criminal records data. In one study criminal activity was measured through incarcerations of 1 night or more (Klein et al., 2012) and in the final study criminal activity was measured during adolescence through parent report (Bor et al., 2004).

The types of behaviours associated with later criminal activity were almost homogenously externalising behaviour types, with 3 studies assessing both conduct and hyperactivity problems, 2 studies assessing hyperactivity problems only, 2 studies assessing conduct and internalising problems and 1 study assessing conduct problems only.

In all eight studies, early behaviour problems of any kind were significantly related to later criminal activity.

In one of the three studies linking both conduct and hyperactivity to criminal activity 305 participants (M=230) were followed from age 9 years to the age of 26. Childhood behaviour was measured using parent and teacher ratings on standardised scales and criminal activity was measured through both self-report and criminal records data (Babinksi et al., 1999). The results revealed significant effects of conduct problems (OR = 7, p= .001) and hyperactivity (OR 2.85 p =.015) in males but not females. Similar results were found in the other 2 studies.

Using a slightly larger sample (N=828) and following children from a younger age (3 years) Stevenson and Goodman (2001) also reported a significant relationship between parent rated behaviour problems and later criminal activity at age 23-24 years using criminal records data.

Both hyperactivity and conduct problems during childhood were again shown to increase the chances of offending whereby hyperactive children were almost twice as likely to commit an offence (OR 1.86 p<.01) and children with conduct difficulties were also almost twice as likely to commit an offence (OR 1.91 p<.01).

These findings were further replicated by the final study assessing both conduct and hyperactivity problems (Simonoff et al., 2004). In this study 225 participants aged 28-59 (mean 38.2 years) were asked to retrospectively report whether they were diagnosed with hyperactivity and/or conduct problems before they were 7 years old using the Retrospective child and adolescent psychopathology (RECAP) assessment, a retrospective tool which has been validated on both a general and clinical sample (Holmshaw and Simonoff, 1996).

Associations were then made between these reports and criminal convictions using criminal records data. Univariate analyses showed an association between offending in adulthood and both child conduct disorder (OR =3.36 p<.01) and hyperactivity disorder (OR = 2.32 p<.05).

Interestingly, Simonoff and colleagues also looked at juvenile offending rates that occur below the age of 17. These held much stronger associations with early behavioural problems with much greater likelihoods of committing an offense being associated with conduct disorder (OR = 11.52 p<.001) and hyperactivity disorder (OR = 3.13 p<.001) suggesting the effects may dissipate slightly across adulthood.

Although these results would suggest a clear relationship between early problems and later criminal activity, none of these studies was rated above moderate quality. Only the study by Stevenson and Goodman considered the effects of confounders in the analysis, though Simonoff et al. did account for IQ. This could mean that effect sizes may appear greater than they actually are. Comparing the odds ratios between the three studies would suggest this to be the case with both Simonoff et al. and Babinski et al. both presenting higher values compared to Stevenson and Goodman. It is thus likely that these values are more reflective of the actual effects of early behaviour problems on later criminal activity.

The two studies examining hyperactivity only reported criminal activity in terms of arrest rates and/or incarceration rates. These were strikingly different between the two studies.

Satterfield and Schell (1997) reported arrest rates of 21% for men who had been diagnosed with hyperactivity as a child versus 1% of controls (p<.001), and incarceration rates of 12% versus 0% (p<.01). Klein et al. (2012) on the other hand report incarceration rates for men who were hyperactive children nearly 3 times higher at 36.3%. The number of controls were reporting incarcerations was also considerably higher at 11.8%, though the difference between groups was still significant (p<.001). This difference may be due to large differences in the length of follow up, as the men in Klein and colleagues study were followed to age 41 years compared to only age 25 years in the comparative study. This allows a greater duration of time in which participants may commit an offense. The samples and study quality were otherwise highly similar.

The final three studies provide further support for the link between early conduct problems and later criminal activity.

Firstly, conduct problems at age 5 years measured though parent reports were linked to adolescent antisocial behaviour at age 14 in a large sample of 5278 children as part of the Mater University Study of Pregnancy, a large Australian cohort study (Bor et al., 2004). After controlling for potential confounders children with early conduct problems were significantly more likely to engage in antisocial behaviour according to parent reports compared to controls (OR 3.1 p<.01). Furthermore attention problems were also shown to be significant predictors of antisocial behaviour in adolescence (OR 2.51 p<.01) giving further weight to the importance of externalising behaviours during childhood. Although these findings do not examine criminal activity in the same way as the others, they are still important to note as adolescent behaviour has been shown to be a significant predictor of adult criminal activity (Satterfield and Schell, 1997, Simonoff et al., 2004).

Early conduct problems were also linked to adult criminal activity in terms of rates of offending (Fergusson et al., 2005) and the increased likelihood of conviction (Murray et al., 2010). Both of these studies used patient reports of behaviour before the children were 10 years old. Fergusson et al. categorise children according the percentiles, and show that rates of offending drastically increase with the severity of childhood behaviour symptoms. For example, compared the lowest scoring 50% of children where 4.5% went on to receive a conviction by age 25, 19.5% of the highest score 5% reported convictions, over 4 times as many. This presents an interesting finding showing that it is not just the presence of

behaviour problems during childhood that is important but also the severity of symptoms. No other study was found that split children by severity of behaviour problems in this way. In contrast, Murray et al. assess outcomes in terms of gender differences in the likelihood of children with conduct problems being convicted by age 34 years compared to controls. After controlling for confounders, fairly similar results were found for both genders. Boys with conduct problems were 1.6 times more likely than control boys to be convicted of an offense (p<.05). Girls with conduct problems were twice as likely to be convicted of an offense compared to control girls (p<.05) suggesting that the presence of early conduct problems in girls may be more predictive of future criminal activity than for boys, in opposition of Babinski et al. (1999) which found child conduct problems predictive of boys criminal activity but not girls. Because the rates of criminal activity are generally lower for men than for women in the general population (Stevenson and Goodman, 2001) a larger sample size may be necessary to detect the effects of early behaviour problems in girls. It is thus possible that the relatively small sample size in the Babinski study was insufficient to detect the differences for girls.

Collectively these studies support a strong link between early child behaviour and later criminal activity with conduct and hyperactivity (externalising) problems being especially predictive.

6.3.2.1 Assumptions for the decision model: Criminal activity

As with the educational outcomes, due to large differences in the methods used to assess criminal activity, meta-analysis of study findings was not possible.

One high quality (Murray et al., 2010) conducted in a UK context was identified for inclusion in the decision model. In this study, estimates of the effect of early behaviour problems on criminal activity were stratified by gender, thus an average of the two odds ratios will need to be derived for the model which accounts for all children.

6.3.3 Early childhood behaviour and labour market activity: employment and earnings

Table 6.3. An overview of the included studies including quality assessment rating relating to labour market activity and earnings

Author, year, title	Aim	Sample	Measures	Main findings	Quality
Knapp et al. (2011) Economic outcomes in adulthood and their associations with antisocial conduct, attention deficit and anxiety problems in childhood.	To quantify the connections between childhood behavioural problems and economic based outcomes in adulthood.	Data were used from the BCS. N=17,196	**Age 10** Behaviour: Teacher Rating Scale and the Rutter Teaching Scale. **Age 30** Labour market: self-reported income and employment.	Labour market activity: Conduct (men): B = -.089, p<.05 Attention problems (men): B = -.118 p<.001 Attention problems (women): B= -.111 in women p<.001 Earnings: Conduct problems: Predicted earnings at 25th percentile £499 for men £279 for women, 90th percentile £524 for men £293 for women, 95th percentile £530 for men, £296 for women). Attention: 25th percentile £499 for men £279 for women, 90th percentile £ £463 for men £217 for women, 95th percentile £455 for men £208 for women.	Strong
Caspi et al. (1998) Early failure in the labour market: Childhood and adolescent predictors of unemployment in the transition to adulthood.	To investigate childhood and adolescent predictors of youth unemployment.	New Zealand sample of 954 from Dunedin Multi-disciplinary Health and Development Survey.	**Age 7-9** Behaviour: parent and Rutter Child Scales. **Age 21** Labour market: self-reported periods of unemployment	Multivariate analysis: Coefficient .004 (SE .002) p<.05 Percentage change in the probability of unemployment = 7.7% Behaviour problems - increase in the number of unemployed – 1.1 more months	Strong

Table 6.3. An overview of the included studies including quality assessment rating relating to labour market activity and earnings ctd.

Healey et al., (2004) Adult labour market implications of antisocial behaviour in childhood and adolescence: Findings from a UK longitudinal study.	To explore the effect of difference antisocial development pathways on labour market outcomes to age 32.	Sample of 411 males from the Cambridge study of delinquent development (CSDD). Largely Caucasian sample (87%)	**Age 8-10** Behaviour: peer and teacher ratings **Age 32** Labour market: self-reported period of unemployment and earnings	22% of children with behaviour problems alone reported periods >12 months unemployed over the 5 year period of interest compared to 32.4% of children with behaviour problems who were also engaged in criminal activity during adolescence and 11.1% in controls. Mean earnings of problem children £173.05 pw, problem children with criminal record £153.38, control £175.26. (1985 prices) Expected weekly income of men who had behavioural problems and a juvenile criminal record earned on average 68% less than the rest of the sample, equal to around £50 per week.	Weak
Klein et al. (2012) Clinical and functional outcome of childhood attention-deficit/hyperactivity disorder.	To examine whether children identified as ADHD at age 8 have worse educational, occupational and economic outcomes.	Sample of 135 men diagnosed with ADHD at age 8 as well as 136 controls	**Age 8** Behaviour: clinical diagnosis of hyperactivity and attention problems. **Age 41** Labour market: self-reported income	Median salary for men with ADHD was $60,000 compared to $150, 000 in men without ADHD (p<.001).	Moderate

Table 6.3. An overview of the included studies including quality assessment rating relating to labour market activity and earnings ctd.

Woodward and Fergusson (2000) Childhood peer relationship problems and later risks of educational under achievement and unemployment.	To investigate associations between teacher-rated peer relationship problems at age 9 and subsequent risks of education under achievement and employment by age 18.	964 respondents from CHD study in New Zealand.	**Age 8** Behaviour: teacher interviews **Age 18** Labour market: self-reported periods of employment	Children with behavioural problems are 3.3 times more likely to be without employment (p<.0001) at age 18 compared to controls.	Moderate
Fergusson et al. (2005) Show me the child at seven: the consequences of conduct problems in childhood for psychosocial functioning in adulthood.	*Aim:* To examine the long term consequences of childhood conduct problems.	Data from New Zealand cohort study. 1265 participants.	**Age 7-9** Behaviour: parents and teachers Rutter A2 behaviour questionnaire and Connors behaviour questionnaire. **Age 25** Labour market: self-reported periods of unemployment	No significant differences	Moderate

Table 6.3. An overview of the included studies including quality assessment rating relating to labour market activity and earnings ctd.

| Alatupa et al. (2013) Disruptive behaviour in childhood and socio economic position in adulthood: A prospective study over 27 years. | *Aim: T*o examine the relationship between child behaviour and socio-economic position, income and occupational status in adulthood. | 782 cohort members from Young Fins cohort study in Finland (403 female) | **Age 3-9** Behaviour: Unknown measure

Age 30-36 Labour Market: self-report of job type | Low occupational status (manual/unskilled jobs): Hyperactivity: OR 1.21 p=.015 Conduct OR 1.21 p = .017 Social adjustment: OR 0.8 p= .006

Behaviour not significantly related to income. | **Weak** |

Seven studies were identified that explored the relationship between early behaviour problems and later labour market activity. This was mostly described in terms of employment status and average earnings, though one study reported differences in occupational status in terms of whether occupations were low (unskilled, manual) or high (skilled, professional).

All studies reported associations between early behaviour problems and future employment related outcomes, however these varied in degree.

In one study both conduct problems and attention problems at age 10 years were found to significantly predict the probability of being in employment at age 30 (Knapp et al., 2011), though conduct problems were only predictive for men. Results showed that conduct problems reduced the probability of being in employment by 8.9 percentage points (p<.05), and attention problems reduced the probability of being in employment by 11.8 percentage points for men (p<.001) and 11.1 percentage points for women (p<.001).

Four other studies also noted differences in employment according to early behaviour problems. Caspi et al. (1998) report a significant relationship between early behaviour problems and future unemployment with a percentage change of 17.7%. Behaviour problems were also noted to increase the number of months in unemployment by 2.6 additional months.

Similarly Healey et al. (2004) report that 22% of children with behaviour problems report periods of unemployment greater than 12 months over a 5 year period leading up to age 32 years, a rate double that of controls. When early behaviour problems was combined with criminal activity this figure rose to 32.4% reporting long periods of unemployment. Support for this association was further provided by Woodward and Fergusson (2000) who linked early peer/social problems with unemployment showing that 18 year olds who were classified as having these issues were 3.3 times more likely to report periods of unemployment greater than 3 months compared to controls and more recently from Fergusson et al. (2005) who report that 25 year olds who had high levels of behaviour problems at age 7-9 years were 2.6 times more likely to report long periods of unemployment. Cumulatively these three studies suggest that children with general behaviour problems are between twice and three times more likely to have significant periods of unemployment in adulthood compared to controls.

Knapp and colleagues also reported on the differences in earnings associated with behavioural problems, showing differences in effects of different behavioural traits. The probability of being in employment was combined with predicted earnings of employees to derive an estimate of 'expected earnings'.

As expected, both men and women with childhood attention problems received lower expected weekly earnings than those who scored low for behavioural problems (-£44 for men, -£71 for women between the highest scoring 5% and the lowest scoring 25%).

Comparisons related to conduct problems showed the opposite pattern. The expected earnings for the lowest scoring 25th percentile (acting as the control group) were said to be £499 for men and £279 for women. Expected earnings for those in the worst scoring 10% were actually higher than those for controls at £524 for men and £293 for women, and higher still for those scoring in the worst 5% at £530 for men and £296 for women.

This is an unusual finding, which has not been replicated elsewhere, however the study was deemed of strong quality, controlling for a range of important confounders, and using strong cohort data of 11, 182 participants, meaning the results are less likely to be due to chance or analytical error. It could be that other personality traits associated with conduct behaviour lend themselves to certain professions, such as management, a career associated with higher earnings. However, a recent study by Alatupa et al. (2013) found that behaviour problems during childhood (ages 3-9 years) significantly predict the types of jobs individuals have in adulthood (age 30-36 years) with both hyperactivity (OR 1.21 p=.015) and conduct problems (OR 1.21 p = .017) significantly increasing the likelihood of being in an unskilled/manual job. These findings would negate this theory and suggest that children with behaviour problems should go on to have lower earnings than those without.

Other studies have also shown that early behaviour problems often result in lower earnings. In one study, men who were diagnosed with ADHD at age 8 report much lower salaries than men without ADHD, resulting in a median income of $60,000 compared to $150,000 in this small US based sample (n=271). Similar results have also been found in a UK study, where behaviour problems were linked to lower weekly earnings, particularly when associated with juvenile crime (Healey et al., 2004).

6.3.3.1 Assumptions for the decision model: Labour market activity and earnings

One high quality study explored the effect of early behaviour problems on earnings in adulthood within a UK context and was selected for inclusion in the decision model (Knapp et al., 2011).

A general lack of high quality research pertaining to labour market activity was identified with UK based research being particularly weak. Only 1 study considered educational attainment as a covariate (an important predictor of employment) (Woodward and Fergusson, 2000) thus this will be used as the base estimate for the decision model. Given that the sample of this study originates in New Zealand, a culturally similar country to the UK, this study is further justified as a useful indictor for a UK population.

6.4 Estimates for the decision model

This review sought to identify potential literature that could be used to inform a decision model linking early childhood behaviour problems to long-term outcomes. Given the expected breadth of the literature a scoping study design was employed to identify the main types of evidence available and to capture studies using a diverse range of methodologies.

A total of 21 studies were identified that explored these associations. Though the quality of these studies varied, findings showed relatively consistent lasting effects of early behavioural problems on all three longitudinal domains of interest.

Due to the wide range of methodologies meta-analytical techniques could not be applied. Consequently the highest quality studies from each domain were selected to provide estimates of effects for the decision model. UK data were favoured where available. Table 6.4 details the estimates that would be used to inform the decision model.

Table 6.4. Parameters for the decision model

Educational attainment	
Washbrook et al. (2013)	Hyperactivity problems: Boys failure to achieve 5 A*-C: OR 1.57 p<.01 Girls failure to achieve 5 A*-C: OR 1.47 p<.01 Conduct problems: Boys failure to achieve 5 A*-C: OR 1.39 p<.01 Girls failure to achieve 5 A*-C: OR 1.46 p<.01
Criminal activity	
Murray et al., (2010)	Conduct problems: Boys being convicted of an offence: OR 1.6 p<.05 Girls being convicted of an offence: OR 2.0 p<.05 Internalising problems: Boys being convicted of an offence: OR 0.6 p<.05
Labour market activity and earnings	
Woodward and Fergusson (2000)	Peer problems: Likelihood of unemployment at age 18: 3.3 (p<.0001)

Chapter 7: The cost-effectiveness of screening for postnatal depression for child behavioural outcomes

7.1 Introduction

As discussed in Chapter 1, previous research exploring the cost-effectiveness of screening for postnatal depression concluded that screening was not cost-effective (Hewitt et al., 2009). This model had two key limitations acknowledged by the authors. Firstly, it was not possible for the researchers to take into account the effect of false negatives i.e. those women who would not be opportunistically identified by a doctor but who could be identified by a screen, and secondly the effects of maternal depression on children were not included in the model.

This chapter aims to build on the decision model constructed by Hewitt and colleagues to include the effects of failing to diagnose and treat maternal depression on long-term mental health and child behavioural outcomes using the data derived in Chapters 4 and 5.

7.1.1 Findings of the HTA report

As a consequence of the uncertainty and lack of research and data available regarding the cost-effectiveness of routine screening for postnatal depression in primary care, the UK National Institute for Health Research (NIHR) Health Technology Assessment Programme commissioned an extensive review of the clinical and cost effectiveness of formal identification methods for postnatal depression in primary care (Hewitt et al., 2009). This report aimed to provide an overview of methods to identify postnatal depression in primary care and assess their validity, acceptability, and clinical effectiveness through systematic review.

Probabilistic models were developed to estimate, based on available data, the costs and health outcomes for a range of feasible identification strategies. The analysis was conducted from the perspective of the NHS and personal social services, with costs expressed in 2006/2007 prices and health outcomes in terms of quality adjusted life years

(QALYS). As the time scale of the analysis was only 1 year, no discounting of costs and outcomes was applied.

The models evaluated a hypothetical population of women managed in primary care at six weeks postnatal and followed up for one year and consisted of 5 main components, a diagnostic component and 4 mutually exclusive treatment components representing pathways of care, costs and outcomes according to diagnosis group. These could be true positives, false positives, true negatives, and false negatives (Figures 7.1 & 7.2). As discussed in Chapter 1 (section 1.9), NICE recommends guided self-help strategies, home listening visits or structured psychological therapy as treatment options for postnatal depression. A deterministic economic model conducted as part of the NICE guidance showed structured psychological therapy to be less effective but more cost effective than listening home visits where the incremental cost effectiveness ratio of (ICER) of home visits compared to psychological therapy was £9435 per QALY. As a result of this structured psychological therapy was used as the treatment type in the model.
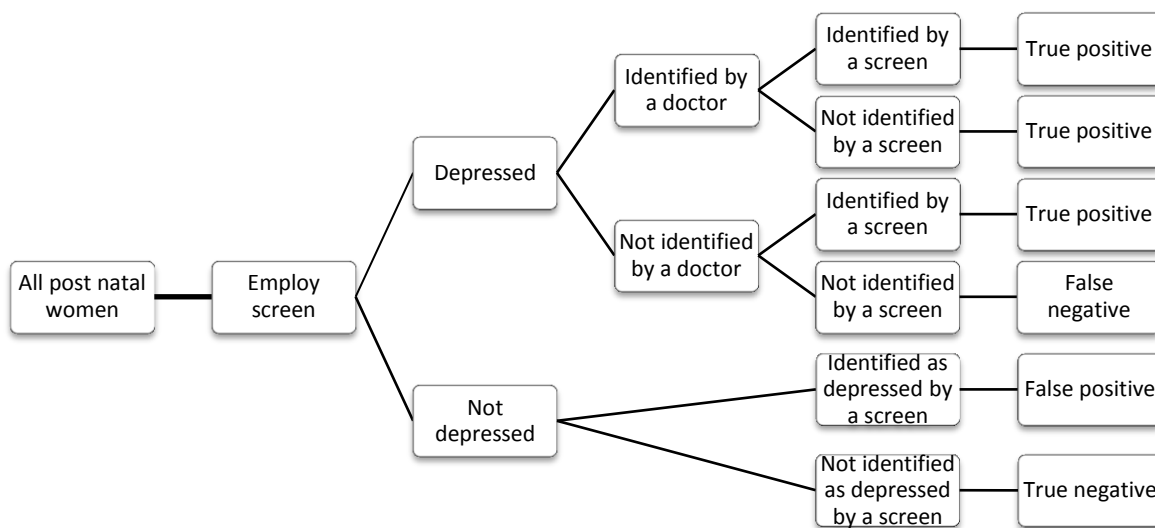

Figure 7.1. The basic structure of the screening model outlined in Hewitt et al. (2009)
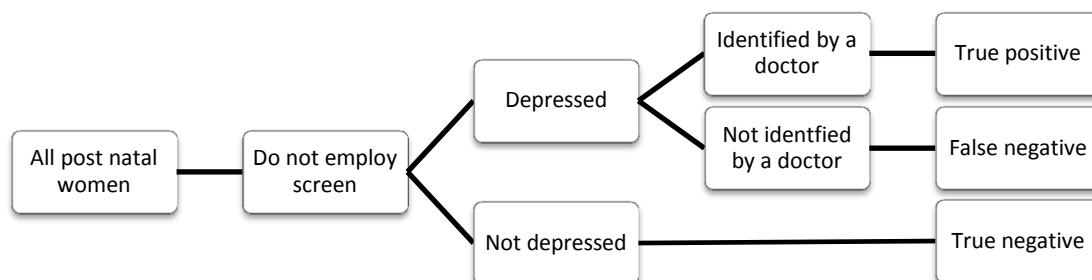
Figure 7.2. The basic structure of the non-screening model outlined in Hewitt et al. (2009)

A range of formal case finding strategies was compared with the current practice of opportunistic case finding with a GP or health visitor. The screening strategies included were the Edinburgh Postnatal Depression Scale (EPDS) and Beck Depression Inventory (BDI) at different cut-off points. One further strategy was considered for inclusion (Hamilton Rating Scale for Depression, HAMD) but was excluded due to insufficient data for inclusion.

Table 7.1 presents the probabilities of various screening tools being cost-effective given a particular willingness to pay for an additional QALY. The standard range of willingness to pay thresholds used by NICE were employed (Rawlins and Culyer, 2004). In the base case analysis, adopting the EPDS with a cut-off point of 16 was found to be the least costly and most effective identification method with an estimated incremental cost effectiveness ratio of £41, 103 per additional QALY compared with routine care alone. The incremental cost effectiveness ratio for all other screening strategies ranged from £49,928 to £272,463 per QALY (Table 7.1). The results showed that no formal identification method seemed to be cost effective under a conventional willingness to pay threshold of £20,000 - £40,000 per QALY meaning that routine care was more cost effective. However, this finding was primarily driven by the costs of false positives. In the base-case analysis the total cost of a false positive was assumed to be £414 which represents the full cost of 'additional care' (consisting of one community psychiatric nurse visit of 1 hour, three GP visits of 10

minutes each and four health visitor home visits of 45 minutes each) considered in the treatment model, but not the treatment itself. To explore the impact of false positives on the results sensitivity analyses were performed using a best case scenario (cost of false positives is £0), and a conservative estimate based on the idea that a 10 minute appointment would be sufficient to correctly diagnose false positive women as not depressed, at a cost of £25.50 per incorrectly diagnosed women. These analyses showed that when less costs were associated with false positives more favourable results of screening were found. Some of the options were ruled out through dominance or extended dominance which in this case is used to eliminate screening tools that have an incremental cost-effectiveness ratio that is higher than that of a more effective tool (Katten and Cowen, 2009). Although routine care remained the least costly and least effective strategy when a false-positive cost of £25.50 was assumed, the next more costly and effective strategy that was not ruled out on dominance grounds was the EPDS cut point 10 (as opposed to cut point 16 in the base-case analysis) whereby the ICER compared with routine care was £29,186 per QALY. The most cost-effective identification strategy was the EPDS cut point 8 at thresholds above £35,390. When the cost of false positives was eradicated, the EPDS with a cut point of 8 dominated or extendedly dominated all other identification strategies and was cost-effective at £25,980 per QALY.

Table 7.1. Base case analysis results (Hewitt et al., 2009, p.103)

Identification option	QALY	Cost	ICER	Probability of cost-effectiveness		
				£20,000	£30,000	£40,000
Routine care	0.846	£49.29	N/A	0.8765	0.5869	0.3930
EPDS 16	0.846	£73.49	£41,103	0.0221	0.0614	0.0684
EPDS 15	0.846	£80.95	ED	0.0070	0.0182	0.0198
EPDS 14	0.847	£94.21	£49,928	0.0158	0.0439	0.0527
EPDS 13	0.847	£110.47	D	0.0052	0.0253	0.0425
EPDS 12	0.847	£109.95	£56,697	0.0177	0.0611	0.0877
BDI 10	0.847	£121.51	D	0.0115	0.0507	0.0895
EPDS 11	0.847	£118.82	£113,411	0.0186	0.0587	0.0853
EPDS 10	0.847	£140.44	£120,968	0.0172	0.0564	0.0890
EPDS 9	0.847	£156.95	£245,210	0.0065	0.0260	0.0464
EPDS 7	0.847	£215.07	D	0.0001	0.0004	0.0012
EPDS 8	0.847	£187.32	£272,463	0.0018	0.0110	0.0245

BDI, Beck Depression Inventory; CE, cost-effective; EPDS, Edinburgh Postnatal Depression Scale; ICER, incremental cost effectiveness
ratio; D represents a strategy ruled out through dominance; ED represents a strategy ruled out through extended dominance.

These results suggest that the cost-effectiveness of screening is largely driven by these false positives. Authors also conducted sensitivity analysis to include a secondary strategy in attempt to reduce the effect of false positives (structured clinical interview). Even with this addition no combined screening strategy was found to be cost-effective.

Because of the lack of sufficient data on the actual costs associated with false positives the authors concluded that the use of formal identification strategies did not appear to represent value for money, based on conventional thresholds of cost-effectiveness used in the NHS.

Although this is a fair assumption based on the data available the model is not without limitations. False negatives were defined in the model as those that had not been identified by either measure. It was assumed that these women would receive usual care leading to; spontaneous recovery, recovery and relapse, or no improvement. It was further assumed that women who did not recover would make an additional appointment with a GP and identification and treatment would possibly follow, though there are likely to be some women who still remain unidentified.

The reduction in the number of doctor-missed false negatives who would be identified by a screening strategy received no consideration in the model. Women who were not identified as depressed by doctors but who were identified by a screen were referred to as true positives. But this is only the case when screening is employed; otherwise they would be false negatives and would not have received intervention. Additionally the model only focused on the costs and outcomes associated with the mother herself, no consideration was given to the costs and outcomes associated with the children of depressed mothers, particularly those not identified, who have been shown to be adversely affected by postnatal depression as described earlier in this thesis (Chapter 4).

The analysis also mapped maternal outcomes over a short period of time (1 year). As discussed in Chapter 5, depression a persistent condition with many women experiencing recurrence of symptoms over time. There is a lack of research considering the differences in maternal depression trajectories over time following an episode of postnatal depression, how these vary according to whether intervention was received, and how this affects child

outcomes. Chapter 5 addresses these questions providing estimates of the effects of different courses of maternal health on child outcomes. The present analysis will use those findings presented in Chapter 5 to develop a long-term decision model based on that presented in the HTA report by Hewitt et al., (2009) to assess whether including the long term effects of false negatives could change the decision regarding the cost-effectiveness of screening for postnatal depression.

The present analysis also addresses a further limitation of the HTA report model which is the exclusive use of QALYs to form the decision. Using QALYs is likely to be an inefficient methodology in this instance as QALYs are both insensitive to mental health (Chrisholm et al., 1997, Shah and Jenkins, 2000) and the important outcomes of postnatal depression extend beyond those which can be represented by QALYs. As discussed in Chapter 6 early child behaviour problems are strongly associated with a range of long-term outcomes including educational attainment, criminal activity and labour market activity and adulthood earnings which are not reflected QALY calculations. These aspects relating to child outcomes should be given consideration in the decision model given the strong links between postnatal depression and child behavioural development. Including such outcomes can be used to assess whether the future benefits relating to intervention can offset of upfront costs.

7.1.2 Economic evaluation: Decision modelling

Economic evaluation has increasingly been used to inform the decisions of health care systems with regards to resource allocation decisions. In the UK the National Institute for Clinical Excellence (NICE) uses economic evaluation to inform decisions concerning medical and diagnostic technologies, surgical procedures and pharmaceuticals (NICE, 2004). As a result of the increasing requisite of economic evaluations to inform decision making, clear requirements have been placed upon researchers in terms of the analytic methods they should be using when conducting economic analysis (Sculpher et al., 2005). Such requirements include the need to incorporate all appropriate evidence into the analysis, to compare new technologies with the full range of relevant alternative options and to reflect uncertainty in evidence in the conclusion of the analysis (Briggs et al., 2006).

The use of decision analysis in health care clinical decision making at both population and individual levels when conditions are uncertain has been increasing and is now commonplace in decision making, often conducted alongside clinical trials (Petrou and Gray, 2011). A decision analytic model uses mathematical relationships to define a series of possible consequences that would result from a set of alternative options being evaluated. Based on the inputs into the model, the likelihood of each consequence is expressed in terms of probabilities. As a result of the growing use of economic evaluation to inform health care decision makers there is increasing importance placed on decision modelling as a vehicle for evaluation (NICE, 2004, Sun and Faunce, 2008).

7.1.3 The Decision Tree

The decision tree is the simplest form of a decision model providing an effective and coherent method for identifying relationships across a broad range of data sets by using a tree-like diagram or model of decisions and their possible consequences. These generally include chance event outcomes, resource costs, and utility consequences. The key features of a decision tree are decision nodes indicating a decision point between alternative options, chance nodes showing a point where two or more alternative events for a patient are possible and reflect the uncertainty of the event, pathways which are mutually exclusive sequences of events and are the routes through the tree and probabilities which show the likelihood of a particular event occurring at a chance node (Briggs et al., 2006).

All decision models are derived through a series of stages. These are: specifying the decision problem, defining the boundaries of the model, structuring a decision model, identifying and synthesising evidence, dealing with uncertainty and heterogeneity and finally, assessing the value of additional research (Petrou and Gray, 2011).

7.1.4 Aim

To develop a decision model exploring the cost-effectiveness of screening for postnatal depression given the effects of false negatives and the effects on children's outcomes.

7.1.5 Objectives

- To examine whether including long-term child behavioural outcome changes the decision regarding postnatal depression screening.
- To examine how accounting for the effects on children of missed maternal depression affects the decision regarding postnatal depression screening.

7.2 Methods

7.2.1 Sample

For this analysis a sub-sample was extracted from the Millennium Cohort Study (MCS). In order to be included in the analysis mothers must have provided complete mental health data at all four data sweeps as well as child SDQ data at age seven. Consequently 6240 mother-child pairings were deemed suitable for inclusion in the analysis. Of the mothers eligible for inclusion 668 were diagnosed and undergoing treatment at sweep 1, 662 were not diagnosed but were exhibiting symptoms of depression on a self-rated scale and 4950 mothers acted as controls showing no indication of depression on any measure.

7.2.2 Measures

7.2.2.1 Maternal mental health

Maternal depression was measured using two different self-assessment tools at different time points. In the MCS every participant was screened for depression when the cohort child was 9 months old (sweep 1) using the Malaise Inventory (Rutter, 1990). This measurement has been described previously in Chapter 4 (section 4.2.3.1). At all further follow up sweeps the Kessler 6 scale was used (Kessler et al. 2002, Kessler et al. 2003) as described in Chapter 5 (section 5.2.5.1). It is important to recognise that these self-assessment tools are not definitive diagnostic tools for case identification of depression but reflect the presence of depressive symptoms. As described in chapter 4, women who were diagnosed and receiving treatment were identified using 2 questions (Chapter 4, section

4.2.3.1). All women who were identified on the self-report measure but not by a doctor are assumed to be false negative cases in this model.

7.2.2.2 Child behavioural development

The SDQ was used to measure the children's risk of having behavioural difficulties at age 7. This assessment has been described elsewhere in this thesis (see Chapter 4, section 4.1.3.2). In the current analysis the dichotomised version of the score will be used in line with the probability requirements of the decision model. This categorises children who are scoring above a recognised cut off point as high risk for having difficulties and those who are scoring below the cut-off point as having low risk of behavioural difficulties. As in Chapter 4 the cut-off point was determined according to the highest scoring 10% of children, giving a cut-off score of 14 points or above. This method for determining cut-off scores has been used in comparable research (Mensah and Kiernan, 2010) and is suggested as an appropriate methodology in the SDQ user guidelines (Youth in Mind, 2010). Based on the original cut-off points determined by Goodman (Goodman, 1997 and 2001), using a cut-off point of 14 and above would identify children with both moderate (14-16) and high (17-40) levels of problematic behaviour.

7.2.3 Group allocation

The mothers' depression status at sweep 1 was used to determine the *'decision'* section of the decision tree. Mothers were first categorised as depressed or not and then further stratified according to diagnosis status, i.e. whether or not they had been diagnosed with depression or were just symptomatic on a self-rated scale. At each subsequent time point mothers were only classified as depressed or not depressed in order to limit the number of possible pathways. At these time points depression status was determined by either measure i.e. diagnosed or self-reported.

7.2.4 The model

Decision modelling was used to map the potential pathways of maternal depression over time derived from four cross-sectional data sweeps taken from the Millennium Cohort Study. Data from the first sweep that took place at 9 months after child birth was used to

derive the initial *decision node*. Chance nodes were added using data from the three available follow up sweeps at 3 years, 5 years and 7 years after child birth. At each chance node women could fall into the 'depressed' or the 'not depressed' option. Two final outcomes were of interest. The first outcome focused on the mothers, specifically, the proportion of women following each pathway. The second outcome of interest was how children's behavioural development was affected by the pathway of depression their mothers followed, measured by whether or not they were classified as high risk for behavioural difficulties on the SDQ at age 7.

The model was used to generate an estimate of the number of children that could be prevented from having problematic behaviour at age 7 through diagnosis and treatment of the mothers postnatal depression. The structure of the model can be seen in Figure 7.3.

Probabilities for children being classified as having behavioural problems were generated by using the 'margins' command in Stata 12 (StataCorp, 2011) using the 17 pathway logistic regression analysis conducted in Chapter 5. This supplies probabilities based on the regression outcome accounting for all covariates whilst being weighted according to MCS sampling weights as described in Chapter 4 (section 4.2.2).

These probabilities are then rolled back (Asleem et al., 2008) to provide a weighted probability of a child developing behavioural difficulties according to whether or not their mother had depression during the first postnatal year that was diagnosed and treated or whether they had a mother who had depression that was not diagnosed and treated accounting for mothers' future depression.

If treatment is found to affect the probability of children having behavioural difficulties, costs and consequences will be incorporated from the scoping review in Chapter 6. This will allow for the exploration of whether future savings could be made through employing early maternal intervention.
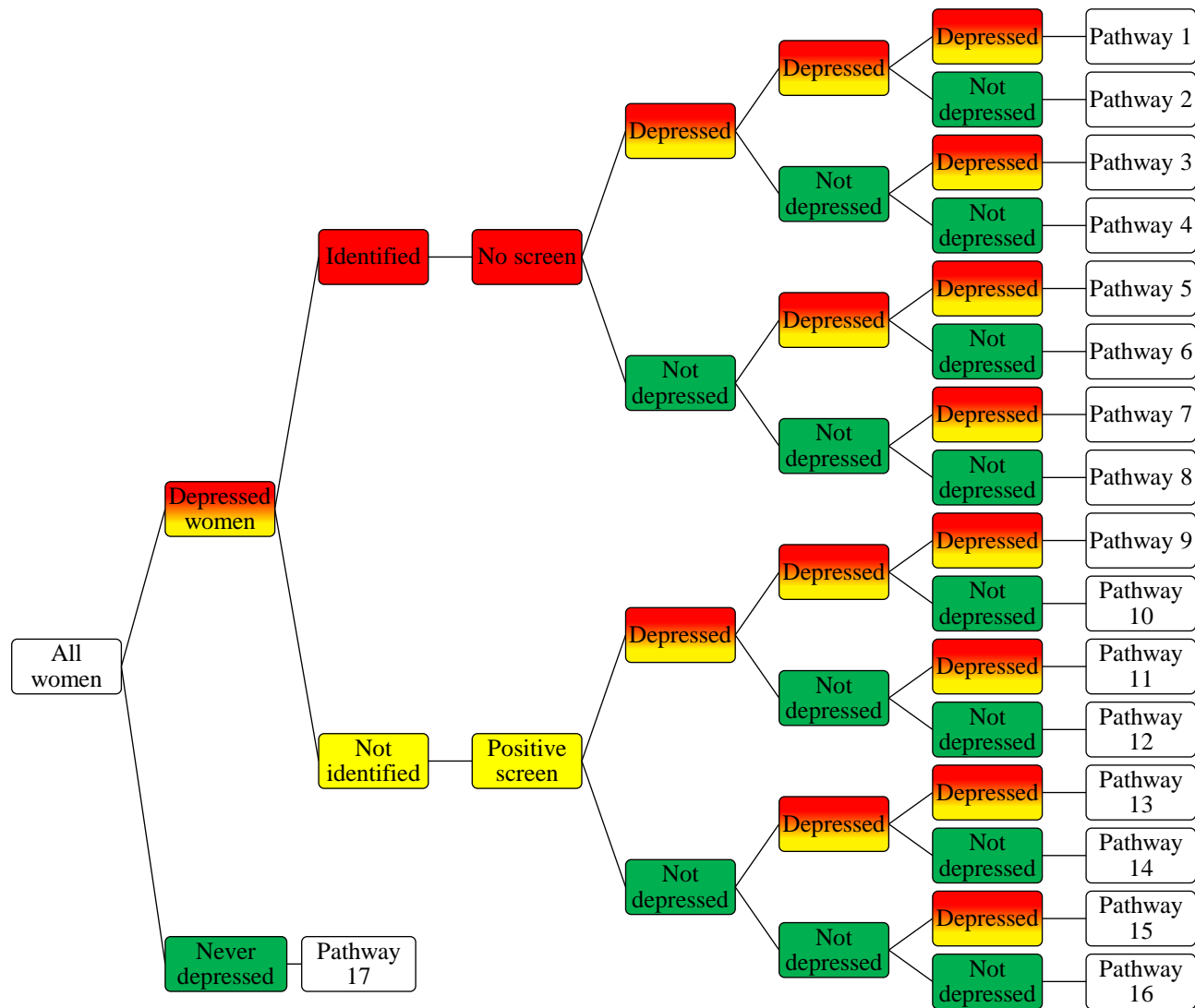
Figure 7.3. Decision model considering the pathways of maternal depression according to whether or not diagnosed and treated at sweep 1

7.3 Results

A probabilistic decision model was generated considering the likelihood of child behaviour problems in relation to maternal depression over time stemming from diagnosis status during the first postnatal year at 9 months (Figure 7.4).

This was produced based on the treatment/identification model by Hewitt et al. (2009) (Figures 7.1, 7.2) but considered recovery, relapse and continuation of maternal depression over a longer timescale and was focused on child outcomes as opposed to maternal symptom relief. Probabilities of children developing behavioural problems were generated for each pathway using logistic regression analysis in STATA 12 (StataCorp, 2011). These were rolled back to give the overall probability of developing behaviour problems according to diagnosis whilst accounting for the future pathways of depression.

The probability of children developing behaviour problems when the mother is diagnosed and treated for depression was shown to be 0.194 compared to a probability of 0.186 when the mother was not diagnosed equating to a difference of -0.008. In order to ascertain a 'real terms' estimate of the difference in the number children this accounts for, the difference between probabilities was multiplied by the total number of mother-infant pairs who were categorised as depressed at sweep 1 (N = 1290). This calculation suggested that when the mother received treatment 10 more children per 1000 displayed signs of behavioural problems at age 7 compared to when the mothers' depression was not identified and treated.

As no benefits were found for identifying and treating maternal depression during the first postnatal year in terms of a reduction in the probability of children having behaviour problems at age 7 there is no likelihood of cost savings. As a result no costing data were added to the model.

224

| | 9 Months | 3 Years | 5 Years | 7 Years | Probability of child behaviour problems |

Figure 7.4. Decision model showing the probabilities of children above threshold for behavioural problems by pathways of maternal depression

7.4 Discussion

The purpose of this analysis was to readdress the question of whether screening for postnatal depression in the population would be cost-effective after incorporating the effects of false negatives (those that would not be identified by a doctor) and the effects of undiagnosed maternal depression on child behavioural development. Previous work by Hewitt and colleagues (Hewitt et al., 2009) used a cost-utility based approach to determine the cost-effectiveness of screening using maternal QALYs. The present analysis aimed to add to this model by taking a cost-benefit approach, weighing intervention costs against future potential gains. Using data from the Millennium Cohort Study across four data collection sweeps from nine months to seven years after birth, probabilities of children scoring high risk for behavioural difficulties at age seven were derived for children of depressed mothers who were either diagnosed or missed by health care professionals during the first postnatal year according to changes in maternal mental health over time.

7.4.1 Principal findings

Using basic probabilistic analysis it was found that screening for maternal postnatal depression using a self-reported assessment is not cost-effective based on findings using the Millennium Cohort Study sample.

After rolling back probabilities generated to estimate the likelihood of children scoring above threshold for behavioural problems at age 7 according to the SDQ, children of mothers who had undiagnosed and untreated depression during the first postnatal year were no more likely to score above threshold than children of mothers who were diagnosed and treated. Children of mothers who received treatment were actually slightly more likely to score above threshold for behaviour problems after accounting for these future pathways having a probability of .194 compared to .186.

Both groups had a much greater probability of developing behaviour problems than children of consistently non-depressed mothers whose probability for scoring above threshold was .067.

Given the lack of a difference in outcomes between the children of diagnosed and undiagnosed mothers a full cost-effectiveness analysis was not carried out.

7.4.2 Relationship with previous findings

This research builds on the findings delivered in the HTA report by Hewitt and colleagues in 2009 by including child outcomes associated with mothers who would not be routinely identified by a doctor.

In this report it was deemed not cost-effective to screen for postnatal depression in the wider population, a decision primarily driven by large costs associated with false positives. The key limitation of this report was the lack of information relating to child outcomes which the present analysis sought to supply.

The findings from the present analysis conform to the decision outlined in the HTA report, maintaining that, at present, it is not cost-effective to screen for postnatal depression even when accounting for future child outcomes, the reason being that intervention for the mother holds no sustainable benefit for the child.

There are two possible explanations as to why no difference was found.

Firstly, it may be the case that the diagnosed mothers experienced more severe symptoms leading them to seek out treatment or to be more readily identified by a doctor, though there is no research exploring this. Although there is limited research in this area what is available has shown that more severe symptoms of depression can have a stronger impact on child development leading to enhanced behaviour and speech problems compared to exposure to mild or moderate maternal depression (Brennan et al., 2000, Hammen and Brennan, 2003).

To explore this idea mean scores were calculated for the mothers following each of the 17 pathways on the Malaise Inventory during the first postnatal year.

Of the 799 diagnosed women who completed the self-report questionnaire, around half of them (48.6% n=388) did not meet the criteria to be classed as depressed (4 points or more).

This could be due to limitations within the measure, the Malaise Inventory scores reasonably well in terms of sensitivity and specificity, 0.90 and 0.65 respectively (Rodgers et al., 1999) meaning that the majority of women with depression would usually be identified by it but the reduced version has not been tested for validity. As we do not know how long the mothers have been receiving treatment for, or what type of treatment they are receiving, it may be the case that these mothers are in recovery at the point of screening. Given that the diagnosed mothers are not indicated as being more severely depressed during the first postnatal year than the mothers in the undiagnosed group this is unlikely to be the reason why children of treated mothers do not do better.

A second and more likely explanation lies in the differences between the longitudinal course of maternal depression between the diagnosed and undiagnosed mothers. Interestingly mothers who had been identified with depression and had received treatment were less likely to fully recover, i.e. not be classified as depressed at any further sweep. Only 17.4 % of mothers who were receiving treatment when their child was 9 months reported no further depression compared to 34.4% in the undiagnosed group. Diagnosed mothers were also more likely to have chronic or recurrent depression (depression reported at all four sweeps) compared to undiagnosed mothers, 29.6% compared to 18.7%. These findings suggest that diagnosed mothers are more likely than undiagnosed mothers to have a longer duration of depression or are more susceptible to relapse which may be due to increased symptoms or an underlying mental health problem. One of the limitations of the MCS is that it does not provide information on history of mental health or for how long the women have been receiving treatment making it impossible to test these differences in this sample.

For both groups full recovery levels are low. This is consistent with other literature documenting maternal mental health after child birth. Although there is little evidence available differentiating between diagnosis based pathways, there is literature to support the idea that women suffer from postnatal depression symptoms in the immediate first postnatal year often then go on to suffer from elevated depressive symptoms at other time points throughout their children's life. For instance McLennan et al. (2001) examined the prevalence, persistence and correlates of depressive symptoms in mothers of toddlers in a nationally representative sample drawn from two linked databases (the 1988 National Maternal and Infant Health Survey and the 1991 Longitudinal Follow up). Depressive

symptoms of 7537 mothers were measured by the Centre for Epidemiological Studies Depression Scale (CES-D) at two time points; when the child was around one year old and again when the child was around three years old. Results indicated that just over a third (36%) reported depression symptoms at both time points which could suggest a high level of persistency or it could be that these were two separate episodes of depression. In a similar study by Horwitz et al. (2007) the number of women reporting a continuation of depression symptoms over time was even greater. In this instance almost half (46.3%) of the women with depressive symptoms when their child was around one year old still had elevated symptoms at a follow up session twelve months later representing a striking consistency in depression symptoms. It is important to note however that both McLennan et al. (2001) and Horwitz et al. (2007) were limited to only one follow up point and neither examined the patterns and persistence of depressive symptoms beyond the first two years of childhood. To address these issues Horowitz et al. (2009) went on to describe the persistence and predictors of elevated depressive symptoms over a slightly longer time frame of 5 years in a sample of 884 predominantly white middle class women. Again over a third of women suffering from postnatal depression (35.6%) reported elevated symptoms of depression on the CES-D at either follow up point at 1 year after birth or 5 years after birth, and 27.4% had elevated symptoms at both follow ups. This study shows that even by the age at which children are beginning to enter formal schooling mothers are still reporting on-going or recurring symptoms of depression. Although the present analysis followed women up 7 years after birth, the findings are in line with these slightly shorter studies, showing high levels of persistency or recurrence of depression in postnatal depressed women.

As discussed in Chapters 2 and 5 multiple or prolonged exposure to maternal depression has greater effects on child outcomes than single exposures. The results in Chapter 5 show that for every additional report of maternal depression the odds of children having behavioural problems at age 7 increased, such findings have also been shown in the wider literature (Brennan et al., 2000, Ashman et al., 2008, Kiernan and Mensah, 2009). Given that women in the diagnosed and treated group are more likely to have on-going or persistent depression, this is likely to be a key driver for the greater effects observed for child behavioural outcomes.

The probabilities of behaviour problems for children of mothers who were identified as depressed at all 4 data sweeps would support this. When mothers were depressed at all 4 sweeps children's probabilities of developing behavioural problems were considerably higher than when mothers were depressed at sweep 1 only (Figure 7.4). This was true for children of both diagnosed and undiagnosed mothers.

It is possible then that the first postnatal year is not a critical period for intervention as future depression may negate any early effects of treatment. Results from Chapter 4 support this idea in showing that although a beneficial effect of treatment is observed until age 7, it is only significant to age 5.

Conversely the first postnatal year could represent a critical period for intervention if the intervention was more successful and efficient at not only treating the maternal symptoms but also in addressing any potential associated parenting problems. As the Millennium Cohort Study was initiated in 2000, before the development of NICE guidelines on depression in 2007, it is likely that the 'optimal' treatment currently suggested by NICE, cognitive behavioural therapy (CBT) was not being provided to the women receiving treatment. Much more likely is it that women were receiving drug therapy through antidepressant medication which has since been recognised in a large gold-standard Cochrane systematic review to be less effective than psychological or psychosocial intervention (Dennis and Stewart, 2004). Given the extensive effects of postnatal depression there is a large potential for cost saving if treatment can be effective for both mothers and their children. More longitudinal research is required to identify and compare appropriate treatment options.

7.4.3 Strengths and limitations

The main strength of this analysis is that it is the first that has derived a probabilistic decision model to determine the probabilities of children developing behaviour problems at age 7 according to diagnosis of maternal depression in the postnatal year whilst accommodating changes in maternal mental health over time.

Previous attempts at evaluating the cost-effectiveness of screening for postnatal depression have failed to include the effects of women who would not be diagnosed routinely without

a screen and also the effects of undiagnosed depression on children (Hewitt et al., 2009). This analysis builds upon previous research to add in the effects of child outcomes associated with undiagnosed maternal depression.

Another key strength of the analysis is that the probabilities generated for child outcomes were drawn from rigorous logistic regression analyses accounting for a range of important covariates as identified in Chapter 3. This means that the probabilities shown are more likely to be accurate and less likely to be due to chance.

A further strength of this analysis arises from the use of a large scale nationally representative sample (MCS) which allows for longitudinal follow up of undiagnosed and untreated depression that would otherwise be unavailable due to ethical implications.

Despite the strengths of the analysis there are also some limitations, one of which is the dichotomised application of the SDQ. As discussed in Chapter 4, there is some concern as to the appropriateness of applying thresholds to SDQ scores. However, there is good research to support that dichotomising between low and higher risk children is a valid methodology (Goodman et al., 2001) and children scoring in the upper 10th percentile have a substantially higher probability of being diagnosed with a psychiatric disorder (OR 15.2 (95% CI 13.0-17.7) and so some confidence can be applied to the results.

At discussed in Chapter 4, one further limitation of the MCS is the self-report scales used to report on maternal mental health are not optimal tools for identifying depression but focus more on 'psychological distress'. In addition to not been geared directly towards postnatal depression the self-report measures are further hindered by the use of different measures at different sweeps. However, both measures have been used in comparative studies (Bartley et al., 2004, Mensah and Kiernan, 2010, Kiernan and Huerta, 2008) exploring maternal mental health, and both have been assessed for validity in providing a good indication of psychological difficulties such as anxiety and depression in community populations (Rodgers et al., 1999, Kessler et al., 2002, 2003). However, it is important to maintain that the validity assessment for the Malaise Inventory is based on the full Inventory which contains 24 items, whereas the inventory included in the MCS adopts only 9 of these questions. This means that the sensitivity and specificity of the 9 item questionnaire may vary slightly from that of the 24 item questionnaire potentially affecting

the rate of false positive results. However, because these screening tools are known to be only indicative of symptoms and cannot provide an actual clinical diagnosis, a number of false positives would be expected when using these measures and caution is applied to the discussion of these findings in light of this. Given that there may be false positives in this sample, the number of women who would be considered to be an undiagnosed missed case of depression (false negatives) in this group may be an overestimate. It is also still likely that some cases are been missed by both measurements.

One further limitation of the study is the reduced number of pathways of depression due to the grouping of women diagnosed and receiving treatment and women with self-reported symptoms at sweeps 2-4. A model with 81 pathways was initially constructed to reflect pathways of depression that fluctuated between being in treatment and being symptomatic without treatment. However this was deemed an overly complex model and resulted in very small numbers of women per pathway diminishing the comparability of child outcomes. As a consequence of this we cannot extract information on differences in outcomes according to future sessions of treatment for depression after sweep 1.

7.4.4 What does this study add?

The findings from the present analysis add new information on the longitudinal course of maternal depression from the first postnatal year according to whether depression was diagnosed or undiagnosed (self-reported) 9 months after birth. Data from the MCS used in this analysis showed few women fully recover following treatment for depression in the first postnatal year (27.4%), with a large number of women maintaining symptoms of depression at each of the follow up time points up to 7 years after birth (29.6%). Though slightly more favourable outcomes are found for undiagnosed women, 34.4% fully recover and 17.5 % of women report depression at all 3 follow up time points, there appears to be a high number of women experiencing consistent or frequent episodes of depression after having failed to be diagnosed and treated.

The analysis also builds upon previous research exploring the cost-effectiveness of screening for postnatal depression in mothers, finding that even when longitudinal child outcomes are incorporated into the model, the decision remains that is not cost-effective to screen for postnatal depression.

7.4.5 Implications and further research

For a program to be considered cost-effective it must first be effective. If screening is to be considered a cost-effective intervention appropriate treatment must also exist. Further research is required to identify the effects of different methods of treatment on both the mothers' outcomes and the children's outcomes.

Given the obvious importance of mother-child attachment as stipulated by attachment theory described in Chapter 2, treatments that include both symptom management and parenting skills are likely to be the most effective.

Given the extent of the costs associated with early child behaviour problems identified in Chapter 6, it is highly important that investment in researching such treatments is made. Not only this, but the findings also show that there is a clear problem with persistent depression, which occurs both when depression was treated and untreated. This implies that current treatment measures may not be sustainable and alternative treatment options should be explored. However, providing adequate treatment for depression is challenging. As discussed in Chapter 5, depression is often a persistent illness with individuals often facing frequent relapses and never experiencing full recovery (Joseffson et al., 2007). This may be particularly problematic when life stressors such as poverty and marital discord which can persist for some time are contributing to depression. This presents a difficult problem for therapists who only have the opportunity to present individuals with coping strategies for stressful circumstances or events and have little influence over such circumstances. It may then be that a more holistic approach to treating mental health is required involving a range of professionals such as education/ employment guidance and social workers.

The MCS data is limited in that mothers were not asked what type of treatment they were receiving so this could not be considered in the present analysis. Because the difference in child outcomes was so small, it would also appear that treatment for the mother does not greatly benefit the children. More research needs to be done comparing treatments that are effective not only for maternal symptoms, but also in preventing developmental problems in children.

7.5 Conclusions

Treatment received by women in the Millennium Cohort Study was not beneficial for children in the long-term meaning that even when considering child outcomes and the vast effects of undiagnosed depression, the decision derived in the cost-effectiveness analysis by Hewitt and colleagues (2009) that it is not cost-effective to screen cannot be challenged. The large effects on child behaviour problems and the long-term effects and costs associated with these mean that child outcomes should be considered in any future analysis should different treatment options be compared.

Treatment appears to have little long-term benefit for the mothers with only a small proportion of mothers fully recovered following intervention and held no long-term benefit for children. Further research is required to identify a more effective treatment approach, only then is screening likely to become cost-effective.

Chapter 8: Discussion

8.1 Introduction

Previous research estimating the cost-effectiveness of screening for maternal postnatal depression has been hindered by a lack of knowledge on the long-term effects on children of undiagnosed depression. Based on the outcomes of screening for the mothers alone over a one year time frame, it was concluded that screening was not cost-effective (Hewitt et al., 2009). However, given the consistently reported negative effects of maternal depression on children, it was hypothesised that this decision may be reversed if child outcomes were available for inclusion in the model. The research carried out in this thesis thus sought to identify the long-term effects of undiagnosed maternal depression on child development and develop a more comprehensive decision model to address the question of whether screening could be beneficial and cost-effective given child outcomes.

8.2 Overview of results

The early years of life are highly important for development (Heckman, 2007, Carneiro, 2003) with a wide range of factors thought to be influential, such as poverty (Kiernan and Mensah, 2009) and parental levels of education (Chaimey et al., 2006). These factors were identified and described early in this thesis in Chapters 1 and 3 and have served as covariates within the analyses to ensure robustness in the results so that effects observed are likely to be true effects associated with maternal mental health.

Many of the factors thought to affect child development can be described as family level characteristics. As discussed in Chapter 1, the family environment plays a powerful role in shaping adult outcomes, and is the grounding of social, cognitive and behavioural development (Heckman, 2006, Ranson, 2008, Ali, 2011).

Although the family environment as a whole is believed to contribute to development, maternal mental health appears to be a predominant factor, consistently being emphasised in the literature as key to development. The relationship between impaired maternal mental

health and child development has been well documented in systematic reviews (Grace et al., 2003, Goodman et al., 2011) and more recent empirical research (Kersten-Alvarez et al., 2012, Agnafors et al., 2013), as discussed in Chapter 1.

Findings from this thesis provide further evidence of this relationship whilst also highlighting the importance of undiagnosed maternal depression in the first postnatal year (Chapter 4), and changes in maternal mental health throughout childhood (Chapter 5), both of which have not previously been explored.

Using these results a comprehensive decision model was created considering maternal depression over time, according to diagnosis during the first postnatal year, and the effects this has on children's behavioural development. A scoping review was also carried out to identify the long term effects of early behavioural problems to identify the longitudinal costs associated with poor childhood behavioural development in terms of educational attainment, criminal activity and labour market activity and earnings.

Despite the effect of undiagnosed maternal depression on children's behaviour, and the substantial longitudinal consequences of this, the decision model showed no evidence that screening for postnatal depression could be cost effective. This conclusion was based on benefits for children of mother-directed treatment that were very small and did not reach statistical significance.

8.3 Comparison with other literature

8.3.1 Importance of maternal mental health diagnosis for child development

As discussed in Chapter 1, maternal mental health is highly important for child development, particularly behavioural development. Findings from the research carried out in this thesis provide further empirical support for this, through both the update of the systematic review carried out in Chapter 1, and the empirical analysis carried out in Chapters 4, 5 and 7.

In addition to providing confirmatory evidence on the importance of maternal mental health, the work carried out in this thesis also identifies the importance of undiagnosed maternal depression, an undeveloped area of research. Although an association between maternal mental health and child outcomes has been documented using both self-reported depression (Kiernan and Mensah, 2009, Murray et al., 2010) and diagnosed depression (Murray, 2003, Pawlby et al., 2008), measures have not been applied to the same population and studied comparatively. The present research is thus the first to apply this methodology and in doing so has shown that the effects of undiagnosed maternal depression are greater than the effects of diagnosed and treated maternal depression up to age 5, indicating that identification and intervention through treatment may be beneficial, but that these effects may not be sustainable to school age (7 years). This lack of stability appears to be a consequence of greater levels of persistent depression over time within the group of diagnosed women, as indicated in Chapter 5, which appears to have a cumulative effect on child development causing mean behavioural scores to rise over time.

Although there is little previous research exploring the long-term mental health outcomes of women identified with postnatal depression, and the effect this has on children, there is some evidence supporting the cumulative effect of exposure to prolonged/ recurrent maternal mental health on children (Brennan et al., 2000, Ashman et al., 2008, Kiernan and Mensah, 2009) thus supporting the results of Chapter 5. In addition to being in line with previous literature, the findings from Chapter 5 also provide further support for the life course approach, a theoretical perspective (described in Chapter 2) emphasising the potential cumulative effects of repeated or prolonged exposure to negative stimuli/ circumstances on development.

Also in line with previous literature is the finding that treatment for diagnosed postnatal depression does not have lasting benefits for children. This relationship has been previously reviewed (Poobalan et al., 2007), and treatment was consistently found to be ineffective at improving the mother-child relationship, a key force in shaping development. Although the analysis conducted in this thesis could not make assumptions about the mother-child interaction due to limitations in the data set, the overall results from this thesis provide support for the ineffectiveness of current treatments of postnatal depression in benefiting child outcomes. As the treatment types received by women in the MCS are

unknown, it is not possible to fully compare the findings of the Poobalan review with the present work.

8.3.2 Cost-effectiveness of screening

The research carried out in this thesis extends a previous decision model in addressing the cost effectiveness of screening for postnatal depression (Hewitt et al., 2009) to include the long-term consequences of undiagnosed maternal depression for children which were not previous accounted for. The importance of undiagnosed depression, expressed as false negatives in the model, is emphasised alongside the importance of accounting for child outcomes, but the recommendation not to screen is unchanged.

The findings of this research also suggest that decision modelling should take into consideration long-term outcomes, as maternal depression often persists well beyond the one year measured in the HTA report.

8.4 Interpretation of findings

As discussed in Chapters 2 and 4, differences in parenting style and the quality of the mother-child attachment are thought to be the likely mechanism through which maternal mental health impacts on children (Field, 2010, Murray et al., 2003, Milgrom et al., 2004). Attachment theory has an evolutionary basis and is centred on the notion that attachments are imperative for survival and thought to be biologically programmed (Thompson, 2006). Consequences of inadequate attachment are thought to extend into all aspects of development, particularly behaviour and social interactions (Cicchetti et al., 2006, Carlson et al., 2004, Goldwyn et al., 2000). Although the concept of attachment and its role in the relationship between maternal mental health and child outcomes could not be explored in this thesis, no competing theory has emerged as a consequence of this research, thus attachment theory remains the most likely explanation of this effect.

Although the results could not provide evidence for attachment theory, they are able to address the two main principles of the life course approach, the critical periods model and the accumulation model.

The early years, particularly the first postnatal year, have been suggested as a critical period for development (Barker, 1998, Kuh et al., 2003). As shown in Chapters 4 and 5, exposure to impaired maternal mental health during this period has lasting effects throughout childhood, even when the exposure is isolated and followed by consistently good mental health demonstrated through no further evidence of depression after the first postnatal year over the course of follow up. This was true for children exposed to either diagnosed or undiagnosed maternal depression with both groups performing significantly worse on behavioural measures than children of non-depressed mothers. This emphasises the enduring effects of exposure during the theorised critical period for development, with effects persisting beyond school onset when the opportunity for problematic behaviour correction is likely to be higher. Such findings have been replicated in previous literature (Mensah and Kiernan, 2010).

Although these findings show that exposure during the first postnatal year alone is sufficient to shape long-term development, thus highlighting the importance of this period for development, the postnatal year may not be a specific and exclusive critical period where exposure to maternal depression will have an effect. It may be the case that exposure to depression at a later time point during childhood can also affect development, which would suggest that maternal mental health throughout childhood, not just the initial year after birth is important. Findings presented in the review by Goodman et al., (2011) discussed in Chapter 1, would support this theory. Although for the purpose of this thesis the review was only updated in terms of the literature focused specifically on depression in the first postnatal year in line with the overarching research questions, the review by Goodman and colleagues extends this time frame and notes the importance of maternal mental health across childhood. Further modelling is required to explore the differential effects of first exposure to impaired maternal mental health at different time points throughout childhood. Whilst this is an important avenue for future research, this question was not addressed in this thesis as it does not add insight into answering the question of whether mothers should be screened for postnatal depression and how undiagnosed maternal depression during this particular period affects children longitudinally. It may be that exposure at any time point in childhood equally impacts children's development, however, the first postnatal year and the increased contact between the mother and health care professionals presents a unique opportunity for identification of depression that is

currently unavailable in the later stages of childhood, thus making it an optimal time for intervention to take place.

In addition to supporting the critical periods model, the analyses conducted in Chapter 5 also yield support for the other model of development of the life course approach to development, the accumulation model, which emphasises the importance of experiences over time and the additive effect these can have on development (Kuh et al., 2003).

In separating out maternal mental health pathways over time by taking a theoretically driven decomposition modelling approach, statistical models were generated to explore the effects of cumulative exposure to maternal depression over time. Collectively, these models clearly showed the additive effect of multiple exposures to impaired maternal mental health to the likelihood of behaviour problems increasing with every additional exposure. The most compelling evidence for this is presented in Table 5.16 (Chapter 5, section 5.3.4), which shows the results of the model assessing the incremental effect of exposure to maternal depression for each additional time point it was reported, regardless of the timing and sequence of exposure, for example whether exposure was at age 9 months, 3 years and 5 years, or at 9 months, 3 years and 7 years, each pathway would reflect '3' exposures.

This research adds to only a small number of studies exploring the effect of maternal depression over time on child outcomes, all of which have converged on the same finding that more persistent and/or recurring depression has a greater effect on child developmental outcomes (Brennan et al., 2000, Ashman et al., 2008, Kiernan and Mensah, 2009). Intuitively, it would be expected that where a mother was identified and treated for depression, as opposed to having depression that is not identified and treated, children would receive some benefit and would not be as severely affected as those children whose mothers who were not identified. This could be due to faster symptom relief, reducing the duration of exposure, or to conscious changes in parenting behaviour the mother may make after being made aware that her depression symptoms are not part of the norm.

In the decision model, statistically derived probabilities of children scoring above threshold for behavioural problems showed that children of diagnosed mothers were actually more

likely to develop behavioural problems at age 7, raising the question of why this might be. As there has been no previous comparison between the outcomes of children of undiagnosed and diagnosed depressed mothers, no explanations currently exist as to why beneficial effects of treatment are not reflected in children's behavioural development over time, only to 5 years. One possible explanation for this is differences in maternal mental health over time between diagnosed and undiagnosed mothers. In the present sample, mothers who were receiving treatment for depression during the first postnatal year were actually more likely to have further episodes of depression than undiagnosed mothers, with 72.6% of diagnosed mothers and 65.6% of undiagnosed mothers reporting depression on at least 1 further time point, and 11% more diagnosed mothers reported the most persistent depression at all 4 time points than undiagnosed mothers. As no history was available concerning mothers' mental health prior to the postnatal data collection, it is possible that diagnosed mothers had greater levels of symptoms or chronic depressive condition. However, scores on the self-reported scale for these women would suggest evidence for lower levels of symptoms, as only 51% of women who were receiving treatment for depression scored above threshold on the Malaise Inventory, implying that these women on average did not have more severe levels of depression at this time point. Additionally the mean Malaise Inventory score for these women was 3.7, slightly lower than the threshold score of 4 required to be classified as 'depressed', and the mean score of the undiagnosed group (mean = 4.9 points). This could be indicative of a treatment effect in terms of symptom relief but this cannot be substantiated without information concerning the history of depression and the type and duration of treatment received. As the mean depression scores showed very little difference between groups, the increased effect on child behaviour problems associated with diagnosed depression is more likely a consequence of the increased likelihood of persistent depression. As discussed in Chapter 5 there is evidence to suggest that women reporting postnatal depression are likely to experience further episodes of depression (Horwitz et al., 2007, Mclennan et al., 2001). This is evident in the current research for both groups.

The fact that children of undiagnosed mothers were more likely to show signs of behavioural problems at age 7 could also be a consequence of inadequacies of the screen, which could have resulted in a high false positive rate, meaning that the women identified as 'depressed' by the screen were not true cases of depression and thus are 'diluting' the undiagnosed group with mothers that should actually be considered non-depressed.

However, although there are likely to be some false positive women in the 'undiagnosed' group, it is evident from the analytical results presented within this thesis that the levels of depression experienced by these women are, on average, sufficient to cause lasting effects in their children's development. Thus despite the fact that these women may not meet the traditional criteria for clinical diagnosis, they may benefit from identification by a medical professional. It thus may be the case that traditional assessments of what constitutes 'depression' may be insufficient for diagnosing new mothers and women who identify as experiencing some symptoms, even those 'too mild' to be considered as 'depression' could benefit from some form of intervention that would ensure limited detrimental effect on their children, should a suitable treatment be available.

It is possible that this issue may have influenced the recommendation of the HTA report. The conclusion that screening is not cost-effective was driven by the high costs associated with false positive cases. It may be that although some of the women scoring above threshold for depression on the screening measures examined in the report were then dismissed as non-depressed during clinical diagnosis procedures, their symptoms could be sufficient as to impact on the quality of their parenting and affect their children.

Despite various types of treatment showing evidence of being beneficial for mothers (Hendrick, 2003, NICE, 2007, McCrone et al., 2004), the results presented in this thesis show that current treatments are ineffective at sustainably benefiting children, showing only short term positive effects. This, combined with the evidence presented in the HTA report which concluded that screening is not cost-effective based on maternal outcomes alone (due to high costs associated with false positives), means that screening cannot be deemed cost-effective and a full economic model was not necessary to readdress the question.

It is possible that the ineffectiveness of treatment is due to the type of treatment received by the MCS cohort mothers. In 2000, when the cohort children were born, NICE guidelines were not yet established regarding the recommended treatment for depression. It is thus likely that women may have received treatments not supported by empirical research such as antidepressants not in conjunction with psychological therapy which have shown limited effects (Dennis and Stewart, 2004). However, even treatments that have been recently established as effective at improving maternal symptoms such as

psychosocial and psychological interventions like CBT or counselling (Dennis and Hodbett, 2007) have proven ineffective at benefiting children, failing to provide significant improvements to the mother-child interaction (Murray et al., 2003, Poobalan et al., 2007), a key factor thought to underlie the relationship between maternal symptoms and child outcomes. The findings of the research conducted in this thesis are therefore relevant to the present context and emphasise the need for more effective treatments.

8.5 What could make screening cost-effective?

8.5.1 Identifying an appropriate screening tool

To increase the likelihood of screening for postnatal depression in mothers being deemed cost-effective several barriers must first be overcome.

A recurring problem across all measures of self-reported mental health is a high false positive rate driven by low levels of specificity, a key factor driving the decision that screening is not cost-effective (Hewitt et al., 2009). This can be reduced by increasing the threshold that women would be required to score to be considered depressed. Indeed when this was explored in the HTA report a score of 16 on the EDPS compared to the standard score of 12 was the closest to being considered cost-effective, ICER £41,103.

Given that the sensitivity of the Malaise Inventory used in the MCS is reported to be 0.90 and the specificity is 0.65 (Rodgers et al., 1999), there is likely to be a number of cases which would be considered by professionals as false positives and thus would receive no further intervention. However, as discussed above, the results presented in this thesis show that self-reported depression is sufficient to have lasting effects on children and thus the criteria to receive treatment may need to be addressed.

In addition to being effective at identifying cases, a screening tool should also be acceptable to the population it serves. This was addressed within the HTA report through means of a systematic review of qualitative literature (Hewitt et al., 2009, Chapter 6).

The overall findings suggested lower acceptability of pen and paper questionnaires compared to screening tools that involved dialogue, though health professionals noted questionnaires were useful to facilitate the discussion of mental health. Across the breadth of the studies women expressed concerns over the repercussions of their answers, leading women from 4 of the 7 studies to deliberately lie. Reasons for this were reported as fear of disclosure to others and moreover the fear of losing their baby and being labelled an incompetent mother (Shakespeare et al., 2003, Poole et al., 2006). Women of ethnic minorities were particularly resistant to discussion of mental health, finding questionnaires culturally inappropriate or difficult to answer (Matthey et al., 2005, Alder, 2007).

This suggests that screening tools need to be brief and encourage dialogue between the women and health care professionals, whilst also be tailored to suit individual cultures. Currently NICE recommends the use of brief case finding questions, the Whooley questions, as discussed in Chapter 7. Although these questions have been found to be effective at identifying cases of depression and have been reported to have a 100% sensitivity rate (Mann et al., 2012), they are recommended for use only as a 'filter' to identify probable cases of depression for further assessment through more in depth screening tools such as the EPDS (Hill, 2010). In addition to being insufficient to guide diagnosis alone, the Whooley questions also failed to meet cost-effectiveness thresholds based on tradition levels of willingness to pay in the HTA report (Hewitt et al., 2009), with the EPDS actually being closer to cost-effective than the Whooley questions thus emphasising the need for more effective screening tools.

8.5.2 Identifying an appropriate treatment

In addition to having an effective and acceptable screening tool, an effective and acceptable treatment also needs to be available. In a systematic review of qualitative evidence drawing on mothers views and experiences, Dennis et al., (2006) reported that women were less accepting of pharmacological treatments, due to fear of addiction (Boath et al., 2004), fear of side effects (Chan and Levy, 2004), fear of stigma (Boath et al., 2004) and fear of transmission through breast milk (Chabrol et al., 2002), all of which were found to significantly influence compliance. The most prevalent treatment preferences centred on talking based therapies, with women preferring opportunities to have in depth discussions about their feelings with a supportive health care professional (Holopainen, 2002, Chan

and Levy, 2002, 2004). Alternative treatments included counselling with partner, relaxation techniques, self-care and physical activity. As with screening, treatments preferences were found to vary between cultures with Asian communities engaging more with community centre based 'activities' such as cooking and crafting classes as means of stress relief rather than a formal treatment which may be seen to bring shame to the woman and the family and be perceived as 'weakness' (Nahas et al., 1999, Parvin et al., 2004, Edge et al., 2004).

Interestingly, none of the treatment options mentioned in the review related to mother-child bonding activities or involvement of the child in the treatment plan. Given that treatment of maternal symptoms alone appears to be insufficient at benefiting children, it is likely that a successful treatment plan would incorporate parenting skills or bonding activities into the programme.

At present there is no recommended treatment that has proven beneficial to both mother and child over the long-term.

It should be acknowledged that providing adequate treatment for depression is challenging. Given that depression is often a cyclic long-term condition treatment providers are faced with the challenge of providing an intervention that is both effective at relieving the immediate symptoms and providing the patient with the tools to deal with depression in the future. As discussed in Chapter 7, this may be particularly problematic when life stressors such as poverty and marital discord which can persist for some time are contributing to depression.

8.6 Implications and further research

Several important avenues for future research have been highlighted throughout this thesis, the most prominent being the need for an effective treatment for maternal depression that is capable of not only providing lasting remediation of the maternal symptoms, but also preventing the detrimental impacts on child development associated with impaired maternal mental health during early childhood.

Pharmacological interventions have demonstrated both limited benefits for symptom relief (Hendrick et al., 2000) and are largely deemed unacceptable treatment methods for women (Dennis and Chung-Lee, 2006) with low levels of compliance and uptake. Large scale reviews have indicated that psychological based interventions such as cognitive behavioural therapy (CBT) have shown promise in treating maternal depression (Dennis and Hodnett, 2007) and are deemed more acceptable though little is known about whether they are beneficial for children. Further research exploring the differential effects of varying types of treatment on child outcomes could thus prove beneficial in establishing which treatment options are most appropriate for new mothers.

In addition to the lack of research measuring the benefits of maternal treatment for depression for children, there is also only a limited evidence base measuring the long-term benefits of treatment for women, and identifying which treatments are most effective at preventing recurrence of depression over time, a seemingly common problem in mothers (Josefsson and Sydsjo, 2007). This is a crucial avenue of research given the evidence presented in this thesis showing the cumulative impact on children of repeated or prolonged exposure to maternal depression over time.

Only once effective treatments are established can the question of whether screening for postnatal depression be readdressed. Given the substantial effects of maternal depression for children, it is likely that with an effective treatment, screening for depression may become cost effective in the future.

8.7 Conclusion

This research has explored the relationship between undiagnosed maternal depression during the first postnatal year and long-term child outcomes in a large UK sample. Using these results a decision model was produced to address whether adding child outcomes to an existing model addressing the question of whether screening is cost-effective would affect the outcome.

The results of this thesis demonstrate that there may be a substantial proportion of postnatal women experiencing depression symptoms that are not identified and treated by

health professionals. Despite the results showing a significant and sustained association between undiagnosed maternal depression and increased levels of behaviour problems in children there were insufficient gains achieved through mothers receipt of treatment following identification. This was reflected through similar outcomes for children of both diagnosed and undiagnosed mothers at each of the measured time points. Whilst there was evidence of some early benefits of treatment present at ages 3 and 5, these were not maintained over the long-term to age 7. This is likely to be due to higher rates of persistent depression in the diagnosed mothers which was found to be associated with a greater effect on children, a fact that further emphasises the inadequacies of current treatment measures.

Given the ineffectiveness of maternal treatment to supply lasting benefits to children, it is concluded that adding child outcomes to the existing cost-effectiveness model would not affect the decision that screening for postnatal depression is not currently cost effective.

Future research should aim to identify a suitable treatment that not only brings about lasting relief of maternal depression but also benefits children and ensures a secure mother-child attachment. Only when a suitable treatment is identified is screening likely to be cost effective.

Appendices

Appendix 1: Chapter 1 Systematic review update

Search strategy adapted from Goodman et al., (2011)

1. mother.ab,ti.
2. limit 1 to yr="2009 -Current"
3. maternal.ab,ti.
4. limit 3 to yr="2009 -Current"
5. mom.ab,ti.
6. limit 5 to yr="2009 -Current"
7. 2 or 4 or 6
8. depression.ab,ti.
9. limit 8 to yr="2009 -Current"
10. depressed.ab,ti.
11. limit 10 to yr="2009 -Current"
12. 9 or 11
13. children.ab,ti.
14. limit 13 to yr="2009 -Current"
15. toddler.ab,ti.
16. limit 15 to yr="2009 -Current"
17. boy.ab,ti.
18. limit 17 to yr="2009 -Current"
19. girl.ab,ti.
20. limit 19 to yr="2009 -Current"
21. adolescent.ab,ti.
22. limit 21 to yr="2009 -Current"
23. 14 or 16 or 18 or 20 or 22
24. behavioral problem.ab,ti.
25. limit 24 to yr="2009 -Current"
26. behavioural problem.ab,ti.
27. limit 26 to yr="2009 -Current"
28. internalizing.ab,ti.
29. limit 28 to yr="2009 -Current"
30. internalising.ab,ti.
31. limit 30 to yr="2009 -Current"
32. externalising.ab,ti.
33. limit 32 to yr="2009 -Current"
34. externalizing.ab,ti.
35. limit 34 to yr="2009 -Current"
36. depression.ab,ti.
37. limit 36 to yr="2009 -Current"
38. anxiety.ab,ti.
39. limit 38 to yr="2009 -Current"
40. anxious.ab,ti.
41. limit 40 to yr="2009 -Current"
42. conduct disorder.ab,ti.
43. limit 42 to yr="2009 -Current"
44. hyperactive.ab,ti.

45. limit 44 to yr="2009 -Current"
46. attention deficit.ab,ti.
47. limit 46 to yr="2009 -Current"
48. mental health.ab,ti.
49. limit 48 to yr="2009 -Current"
50. emotional.ab,ti.
51. limit 50 to yr="2009 -Current"
52. psychopathology.ab,ti.
53. limit 52 to yr="2010 -Current"
54. 25 or 27 or 29 or 31 or 33 or 35 or 37 or 39 or 41 or 43 or 45 or 47 or 49 or 51 or 53
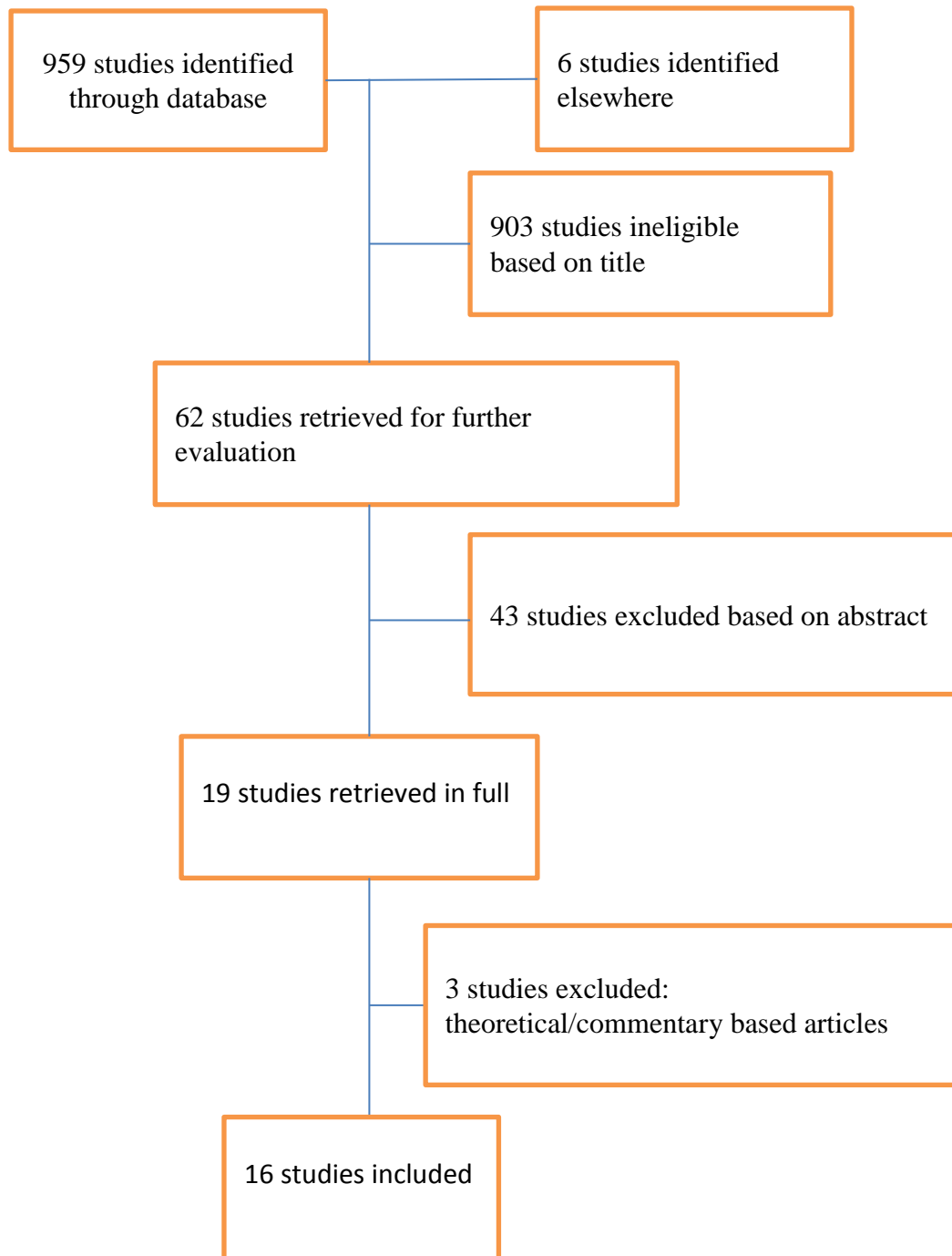55. 7 and 12 and 23 and 54

Figure A1.1 Chapter 1: systematic review update search results flow chart

Table A1.1 Summary of findings from the systematic review update conducted in Chapter 1

Author, year, title	Aim	Sample	Measures	Age at outcome	Main findings	Quality
Agnafors et al., (2013) Effect of gene, environment and maternal depressive symptoms on pre-adolescence behaviour problems – a longitudinal study.	To explore the effects of depression on pre-adolescent behaviour controlling for genetics.	889 pairs from the SESBiC study (Sweden)	Depression: EPDS Child outcomes: CBCL, Hopkins symptom checklist, Life incidence of traumatic events.	12 years	A significant association was found between maternal symptoms of depression and child anxiety and internalising problems OR 5.72 (85% CI: 3.30-9.91).	**Strong**: limited diversity in sample but good size and adequate covariates
Agnafors et al., (2013) Symptoms of depression postpartum and 12 years later associations to child mental health at 12 years of age.	Effect of PND and persistency on child outcomes at age 12.	893 mothers from the South East Sweden Birth Cohort recruited from Child Welfare Centres	Depression: EPDS Child outcomes: CBCL	12 years	Limited association with child externalising behaviour problems Stronger associations with internalising problems – (OR 5.17, p<0.001). Girls were more affected than boys.	**Moderate**: good sample and some covariates
Dietz et al., (2009) Maternal depression, paternal psychopathology, and toddlers' behaviour problems.	Examines the effect of maternal depression on toddler behaviour.	101 mother child pairs from 1 US location. 51 depression, 50 non-depressed	Depression: clinical diagnosis Child outcomes: CBCL (parent rated)	2-3 years	No significant association between maternal depression alone (in the absence of additional paternal psychopathology) and toddler behaviour	**Weak**: Small sample, some covariates
Fihrer et al., (2009) The impact of postnatal and concurrent maternal depression on child behaviour during the early school years	To assess the relationship between PND and child behavioural development	127 mothers and first born infants recruited from parenting project in Australia	Depression: CES-D and diagnostic interview Child outcomes: CBCL and mother interview	6-8 years	PND significantly correlated with children's externalising (r=0.28, p<.01) and children' internalising behaviours (r=0.36, p<.01).	**Weak:** Small sample size, no covariates

Table A1.1 Summary of findings from the systematic review update conducted in Chapter 1 ctd.

Korhonen et al., (2012) A longitudinal study of maternal prenatal, postnatal and concurrent depressive symptoms and adolescent well-being	Effect of PND on adolescent outcomes	Finish study of 349 mother mothers. Current sample based on 191 mothers and 192 adolescents	Depression: EPDS Child outcomes: CBCL and Youth Self Report	16-17 years	Maternal postnatal depressive symptoms were associated with boys' lower Social Competence both in the CBCL (p=0.028) and YSR and Externalizing Problems in the YSR (p=0.012).	**Moderate:** Average sample size, no reference to covariates
Kersten-Alvarez et al., (2012) Early school outcomes for children of postpartum depressed mothers: comparison with a community sample	Examine the effects of PND on school outcomes	142 mother-child pairs from a Netherlands community sample	Depression: Beck depression inventory Child outcomes: Verbal intelligence – Peabody Picture Vocabulary Test-Revised, CBCL	5 years	Children of depressed mothers: Lower ego-resiliency (F = 4.74, p<.05, $n_p^2 = 0.04$) Lower peer social competence (F = 9.48, p<.01, gp 2 = 0.08) Lower school adjustment (F = 4.80, p<.05, gp 2 = 0.04). Girls showed lower verbal intelligence (F= 6.50,p<.05, gp 2 = 0.12) .	**Moderate**: good range of covariates but very small non-diverse sample
Kiernan and Huerta (2008) Economic deprivation, maternal depression	To address the effects of maternal depression on child cognitive and behavioural development	MCS 13, 877 families.	Depression: Malaise Inventory and self-report of clinical diagnosis Child outcomes: Bracken basic concepts scale, SDQ	3 years	Maternal depression statistically effects cognitive, emotional and behavioural development in children.	**Strong:** large sample and large range of covariates

Table A1.1 Summary of findings from the systematic review update conducted in Chapter 1 ctd.

Kiernan and Mensah (2009) Maternal indicators in pregnancy and children's infancy that signal future outcomes for children's development, behaviour and health: evidence from the Millennium Cohort Study.	To identify early life course factors that are associated with children's later development	9249 children from England drawn from MCS	Depression: Malaise Inventory and self-report Child outcomes: Foundation stage profile (teacher assessment), SDQ	5 years	No significant association with educational measures Children were significantly more likely to score above threshold for behaviour problems (OR1.71, p<.001) even after range of covariates adjusted for	**Strong:** large sample and range of covariates
Kiernan and Mensah (2009) Poverty, maternal depression, family status and children's cognitive and behavioural development in early childhood: a longitudinal study	To explore how early circumstances affect child development.	14,777 families from the MCS	Depression: self-report Child outcomes: Bracken Basic Concept Scale , SDQ	3 years	Little association between PND and cognitive outcomes Children were significant more likely to have high levels of behaviour problems compared to peers with non-depressed mothers (F=31.17 p<.001)	**Strong:** sample, good range of covariates
Melchior et al., (2012) Maternal depression, socioeconomic position, and temperament in early childhood: The EDEN mother-child cohort.	Test the hypothesis that depression and socio-economic deprivation contribute to temperamental problems.	Mother- child cohort study in France (n: 1903)	Depression: CES-D Child outcomes: Child temperament - EAS Questionnaire (parent rated)	1 year	Combined maternal depression and family income was associated with children's emotional development (p<.0001) and activity (p=0.02). In multivariate analyses, children growing up exposed to both maternal depression and low income had the highest emotionality scores (β=0.57, se=0.14, p=0.0001).	**Moderate:** good sample and some covariates

Table A1.1 Summary of findings from the systematic review update conducted in Chapter 1

Murray et al., (2010) The effects of maternal postnatal depression and child sex on academic performance at age 16 years: a developmental approach	To explore the effects of PND on academic performance at age 16.	89 participants (50 PND, 39 non-depressed) drawn from an initial community sample of 702.	Depression: EPDS and standardised interview Child outcomes: GCSE grades - G-A* grades were applied a value of 1-8 points to give a total score	16 years	Boys GCSE performance was linked to maternal depression (F = 11.4, p<.01, gp 2 = 0.12) No effects for girls	**Weak:** small sample size, very limited covariates
Murray et al., (2011) Maternal postnatal depression and the development of depression in offspring up to 16 years of age.	To assess child outcomes at age 16 associated with maternal depression	93 (50 index) from Cambridge longitudinal study	Depression: psychological interview Child outcomes: Adolescent mental health measured through diagnostic interviews	16 years	Rate of childhood depression: 41.5% versus 12.5%, OR 4.99 (95% CI 1.68 –14.70).	**Weak:** small sample and limited covariates, sample in only 1 affluent area of UK
Piteo et al., (2012) Does maternal depression predict developmental outcome in 18 month old infants?	To examine the effects of maternal depression during the first 6 months on home environment and child outcomes	360 mother-child pairs from South Australia	Depression: EPDS Child outcomes: Bayley scales and Home Screening Questionnaire	18 months	No significant differences	**Weak:** Average sample, no evidence of covariates, short follow up period

Table A1.1 Summary of findings from the systematic review update conducted in Chapter 1

Quevedo et al., (2011) The impact of maternal post-partum depression on the language development of children at 12 months	To assess the effect of PND on language development	Longitudinal study of 296 mothers in Brazil.	Depression: Mini International Neuropsychiatric Interview Child outcomes: Language segment of the Bayley Scales	1 year	Children of mothers with PND score on average 3 points less on scores of language development p<.01. A longer duration of PND was associated with greater effects.	**Moderate:** average sample size, some covariates
Verbeek et al., (2011) Postpartum depression predicts offspring mental health problems in adolescence independently of parental lifetime psychopathology	To explore the effect of PND on adolescent mental health	2729 adolescents from the TRacking Adolescents' Individual Life Survey in Netherlands	Depression: self-identified cases of PND Child outcomes: CBCL	10-15 years	No significant association with externalising behaviours Significantly higher levels of internalising problems in adolescents whose mother had PND (difference in z-score 0.18, CI 0.04-0.31).	**Moderate:** good sample, some covariates accounted for
Walker et al., (2013) Reported maternal postpartum depression and risk of childhood psychopathology	To explore the effects of PND on emotional and behavioural development	1357 mother-child pairs from the NLSCY in Canada	Depression: self-report of experiencing PND Child outcomes: NLSCY Child behaviour scales which consists of compiled questions from a range of known child behaviour assessment tools such as CBCL	2-5 years	Emotional problems: OR 2.38 (CI 1.15-4.91) Hostile/ineffective parenting: **Age 2-3:** Hyperactivity OR=1.88 (CI 1.14-3.11) Physical Aggression OR = 2.95 (CI 1.77-4.92) **Age 4-5:** Hyperactivity OR = 2.34 (CI 1.22-4.47) Emotional problems OR = 2.16 (CI 1.00-4.67) Conduct Disorder OR = 1.96 (CI 1.09-3.53)	**Strong:** good sample and rigorous covariates

Appendix 2: Chapter 3 Review of reviews

Example of a broad search strategy used in the review of reviews:

1. exp children/
2. Child$.ti.
3. outcomes.ti.
4. development$.ti.
5. review.ti.
6. 1 or 2
7. 3 or 4
8. 5 and 6 and 7
9. remove duplicates from 8

Table A2.1 Summary of findings from studies included in the review of reviews (Chapter 3)

Author	Study findings
Breinholst et al., (2012) CBT for the treatment of child anxiety disorders: A review of why parental involvement has not enhanced outcomes	Unclear whether parent involvement is beneficial for all children. Seems most beneficial when parents themselves have mental health issues which can reinforce child problems. Parental mental health linked to child mental health (anxiety).
Peacock et al., (2013) Effectiveness of home visiting programs on child outcomes: a systematic review	Home visits appear beneficial for all outcomes. Covariates mitigated effects of intervention – where intervention effects weren't found it was attributed to covariates
Orton et al., (2009) Do early intervention programmes improve cognitive and motor outcomes for preterm infants after discharge? A systematic review	Some early intervention effects on cognition but these were not sustained to school age.
Ali (2012) Relationships among Perceived Teachers' and Parental Behavior and Adolescent Outcomes: A Cross-Cultural Literature Review	Quality of relationship with parents and teachers has significant effect on developmental outcomes – school affects contribution of parent-child relationship Unclear where covariates came from.
Deault (2010) A systematic review of parenting in relation to the development of comorbidities and functional impairments in children with attention-deficit/hyperactivity disorder (ADHD)	Few studies explored effects of parenting. Parental mental health and family conflict/oppositional more strongly related with conflict symptoms than hyperactive.
D'Onise et al., (2010) Can preschool improve child health outcomes? A systematic review	Beneficial effects of preschool based interventions on weight, mental health, social competency and crime prevention.
Dewalt and Hink (2009) Health literacy and child health outcomes: a systematic review of the literature	Some relationship between parent literacy and child health care use. Child depression/anxiety problems only found in families where mother had low literacy levels and reported depression.

Table A2.1 Summary of findings from studies included in the review of reviews (Chapter 3)

Jelleyman and Spencer (2008) Residential mobility in childhood and health outcomes: a systematic review	Residential mobility interacts at neighbourhood, family and individual level. Frequent mobility could be marker for risk of behaviour and emotional problems.
Poobalan et al., (2007) Effects of treating postnatal depression on mother-infant interaction and child development: systematic review	Some effects on maternal-infant interaction but few studies were included with generally small samples. Covariates were mentioned in only 1 study.
Thomas and Zimmer-Gembeck (2007) Behavioral outcomes of Parent-Child Interaction Therapy and Triple P-Positive Parenting Program: a review and meta-analysis	Moderate impact of interventions. All cited improvement in child behaviours and parenting stress. Most samples middle class, educated – only 2 studies included lower class, less educated. Not enough to make comparisons.
Chaimay et al., (2006) Risk factors associated with language development problems in childhood--a literature review	Several environmental factors associated with language development
Anderson et al., (2003) The effectiveness of early childhood development programs. A systematic review	Early preschool programs are effective at closing gap between disadvantaged and more advantaged and at preventing developmental delay.
Grace et al., (2003) The effect of postpartum depression on child cognitive development and behavior: a review and critical analysis of the literature	Effects on cognition only evident for boys, mixed evidence for effect on behaviour, SES important for behaviour, 1 study showed boys of depressed and low SES perform worse than any other group
Goodman et al., (2011) Maternal Depression and Child Psychopathology: A Meta-Analytic Review	Maternal depression associated with all forms of behaviour problems. Gender differences more important with older children (>10). Consistent effect of poverty. Lower percentage of married mothers in sample the stronger the effect of depression.

Table 2: AMSTAR is a measurement tool created to assess the methodological quality of systematic reviews.

1. Was an 'a priori' design provided? The research question and inclusion criteria should be established before the conduct of the review.	☐ Yes ☐ No ☐ Can't answer ☐ Not applicable
2. Was there duplicate study selection and data extraction? There should be at least two independent data extractors and a consensus procedure for disagreements should be in place.	☐ Yes ☐ No ☐ Can't answer ☐ Not applicable
3. Was a comprehensive literature search performed? At least two electronic sources should be searched. The report must include years and databases used (e.g. Central, EMBASE, and MEDLINE). Key words and/or MESH terms must be stated and where feasible the search strategy should be provided. All searches should be supplemented by consulting current contents, reviews, textbooks, specialized registers, or experts in the particular field of study, and by reviewing the references in the studies found.	☐ Yes ☐ No ☐ Can't answer ☐ Not applicable
4. Was the status of publication (i.e. grey literature) used as an inclusion criterion? The authors should state that they searched for reports regardless of their publication type. The authors should state whether or not they excluded any reports (from the systematic review), based on their publication status, language etc.	☐ Yes ☐ No ☐ Can't answer ☐ Not applicable
5. Was a list of studies (included and excluded) provided? A list of included and excluded studies should be provided.	☐ Yes ☐ No ☐ Can't answer ☐ Not applicable
6. Were the characteristics of the included studies provided? In an aggregated form such as a table, data from the original studies should be provided on the participants, interventions and outcomes. The ranges of characteristics in all the studies analyzed e.g. age, race, sex, relevant socioeconomic data, disease status, duration, severity, or other diseases should be reported.	☐ Yes ☐ No ☐ Can't answer ☐ Not applicable
7. Was the scientific quality of the included studies assessed and documented? 'A priori' methods of assessment should be provided (e.g., for effectiveness studies if the author(s) chose to include only randomized, double-blind, placebo controlled studies, or allocation concealment as inclusion criteria); for other types of studies alternative items will be relevant.	☐ Yes ☐ No ☐ Can't answer ☐ Not applicable
8. Was the scientific quality of the included studies used appropriately in formulating conclusions? The results of the methodological rigor and scientific quality should be considered in the analysis and the conclusions of the review, and explicitly stated in formulating recommendations.	☐ Yes ☐ No ☐ Can't answer ☐ Not applicable
9. Were the methods used to combine the findings of studies appropriate? For the pooled results, a test should be done to ensure the studies were combinable, to assess their homogeneity (i.e. Chi-squared test for homogeneity, I^2). If heterogeneity exists a random effects model should be used and/or the clinical appropriateness of combining should be taken into consideration (i.e. is it sensible to combine?).	☐ Yes ☐ No ☐ Can't answer ☐ Not applicable
10. Was the likelihood of publication bias assessed? An assessment of publication bias should include a combination of graphical aids (e.g., funnel plot, other available tests) and/or statistical tests (e.g., Egger regression test).	☐ Yes ☐ No ☐ Can't answer ☐ Not applicable
11. Was the conflict of interest stated? Potential sources of support should be clearly acknowledged in both the systematic review and the included studies.	☐ Yes ☐ No ☐ Can't answer ☐ Not applicable

Figure A2.1 AMSTAR tool used to quality assess reviews include in the review (Chapter 3)

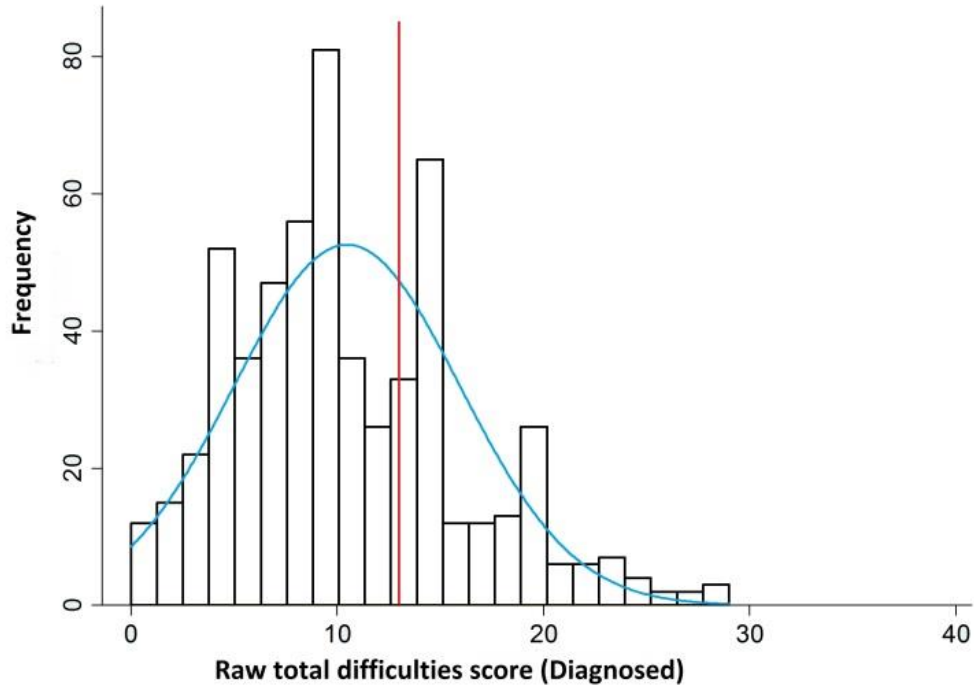# Appendix 3: Chapter 4 Descriptive and diagnostic plots

Figure A3.1 Histogram of the raw total difficulties scores at age 3 years for the children of diagnosed mothers
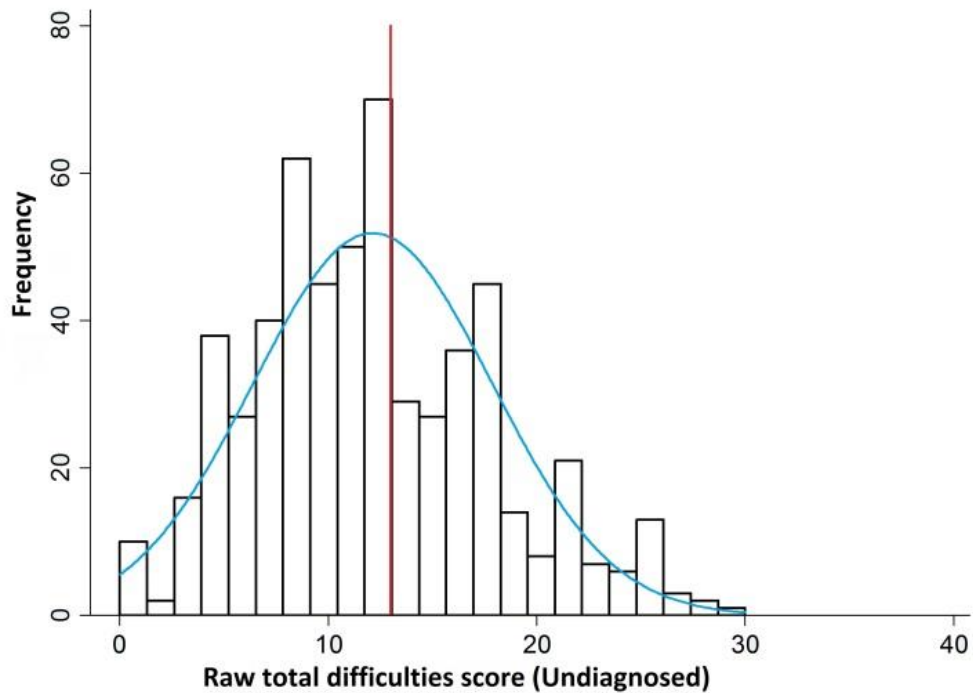

Figure A3.2 Histogram of the raw total difficulties scores at age 3 years for the children of undiagnosed mothers
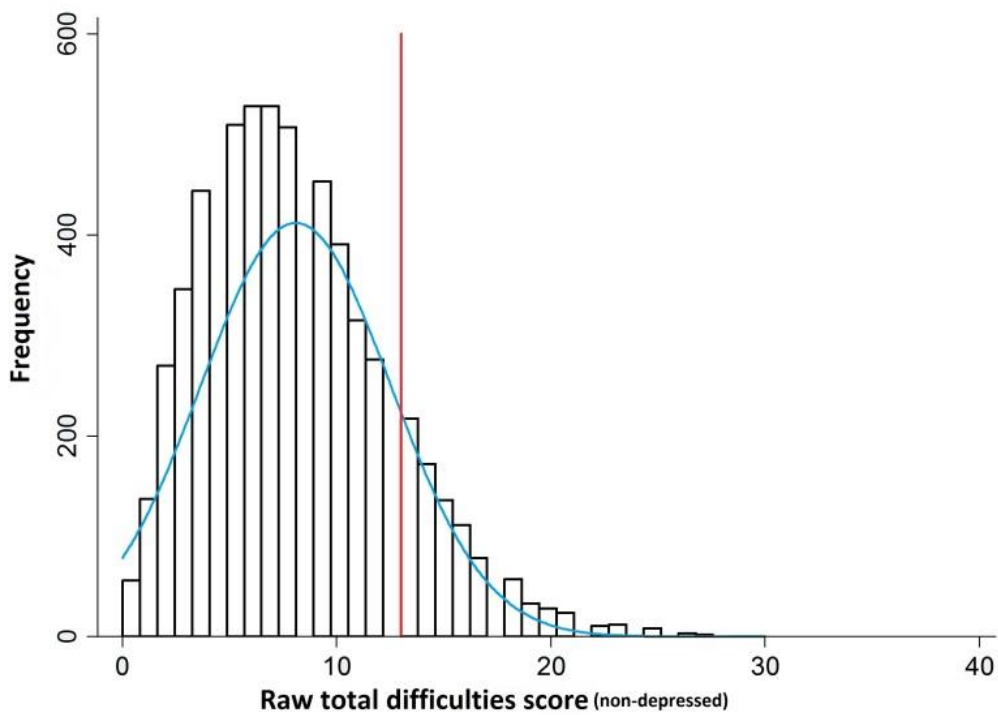
Figure A3.3 Histogram of the raw total difficulties scores at age 3 years for the children of non-depressed mothers

Figure A3.4 Histogram of the raw total difficulties scores at age 3 years for all children

262

Figure A3.5 Histogram of the raw total difficulties scores at age 5 years for the children of diagnosed mothers

Figure A3.6 Histogram of the raw total difficulties scores at age 5 years for the children of undiagnosed mothers

Figure A3.7 Histogram of the raw total difficulties scores at age 5 years for the children of non-depressed mothers
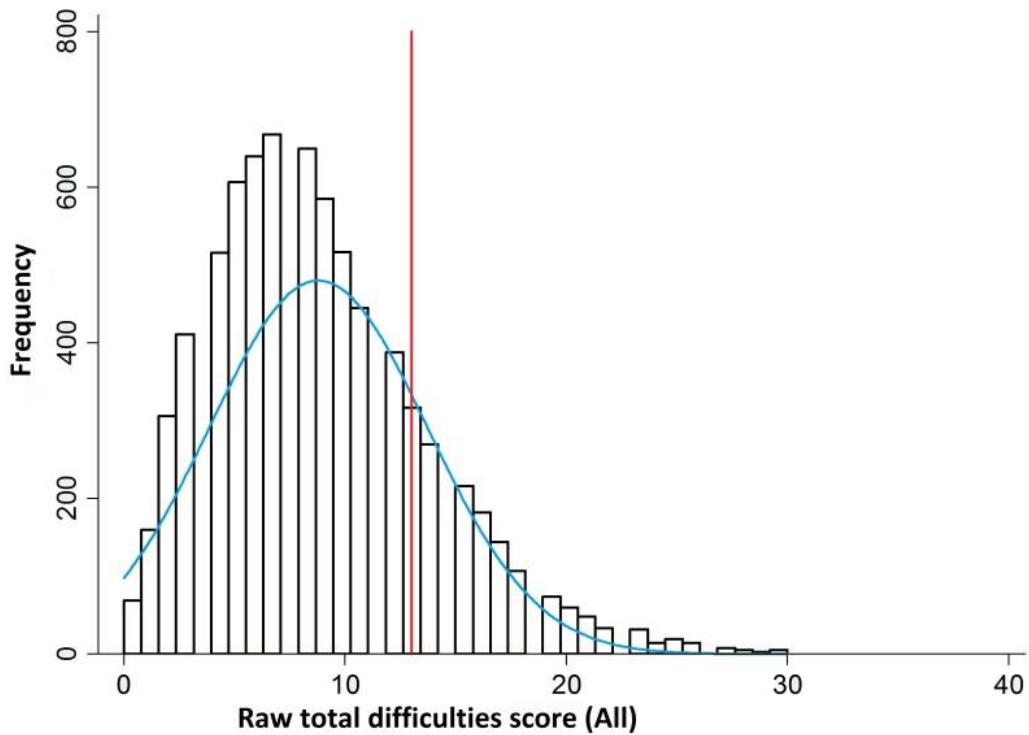

Figure A3.8 Histogram of the raw total difficulties scores at age 5 years for all children
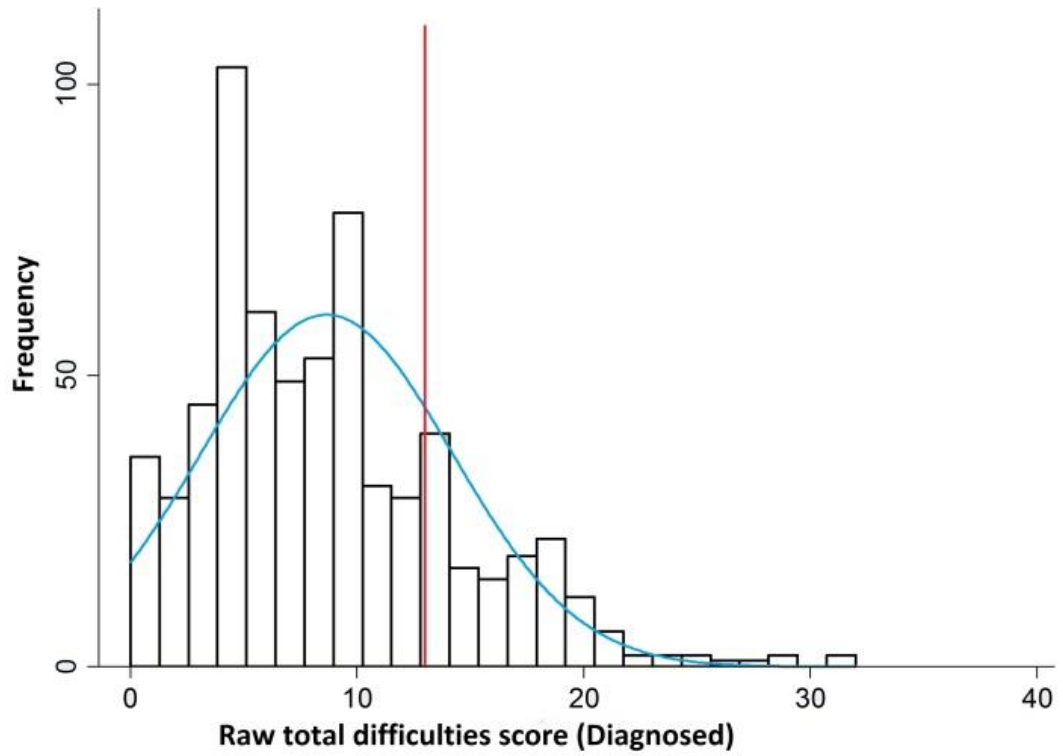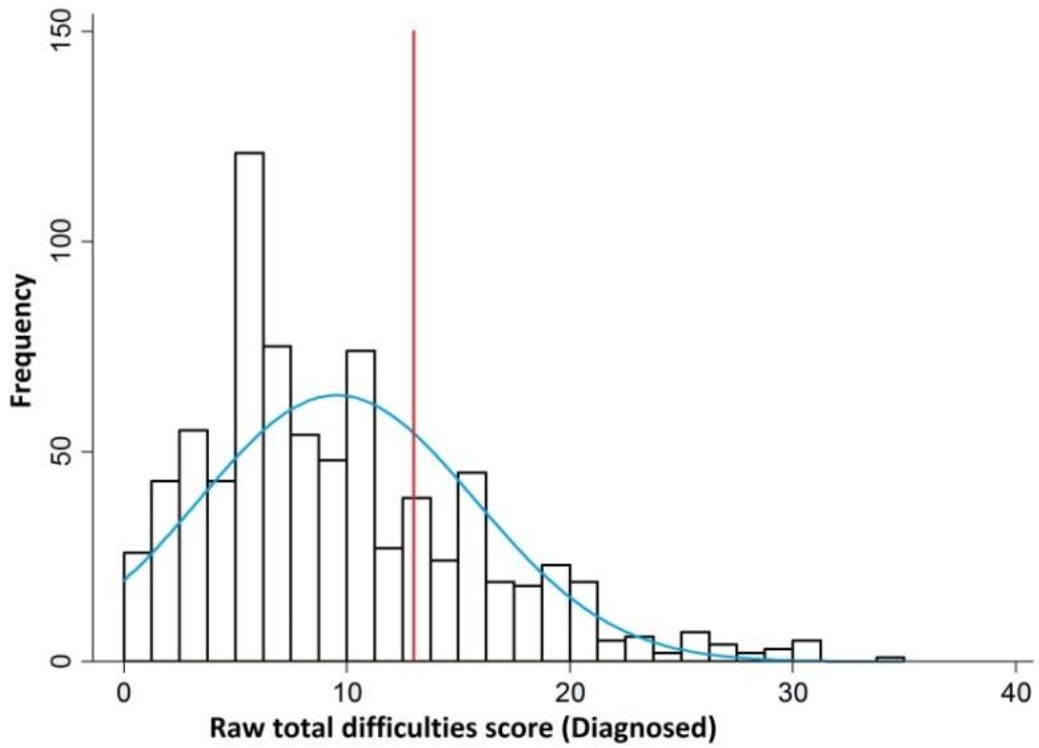
Figure A3.9 Histogram of the raw total difficulties scores at age 7 years for the children of diagnosed mothers
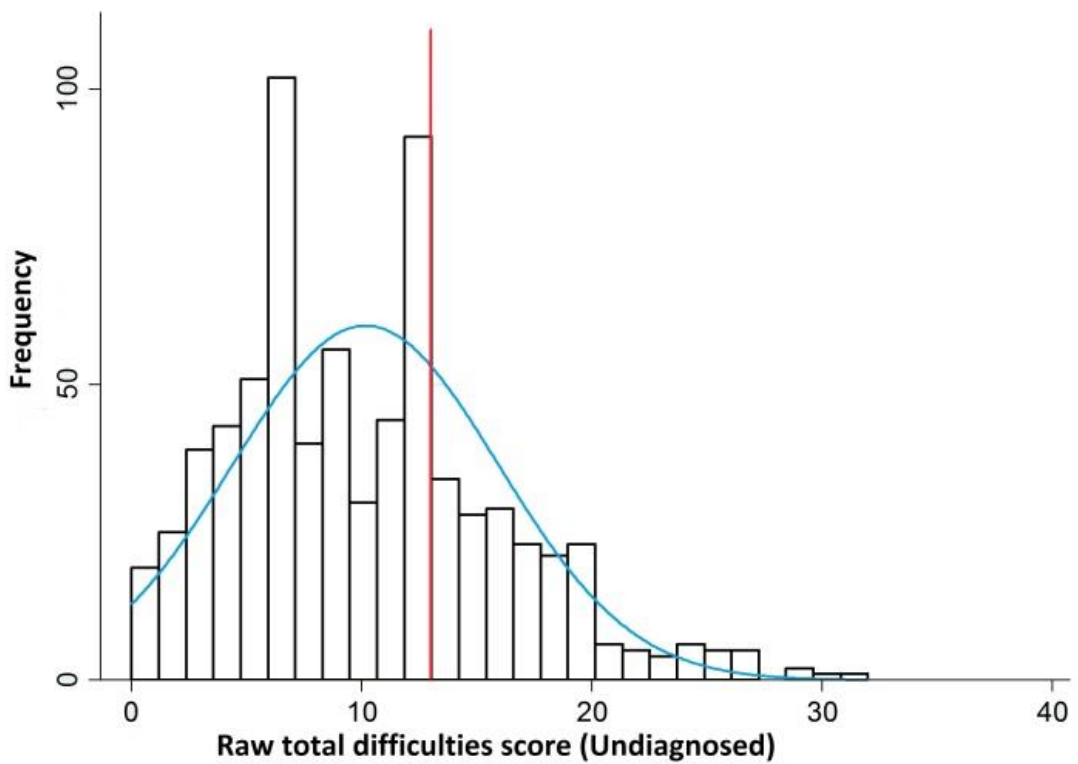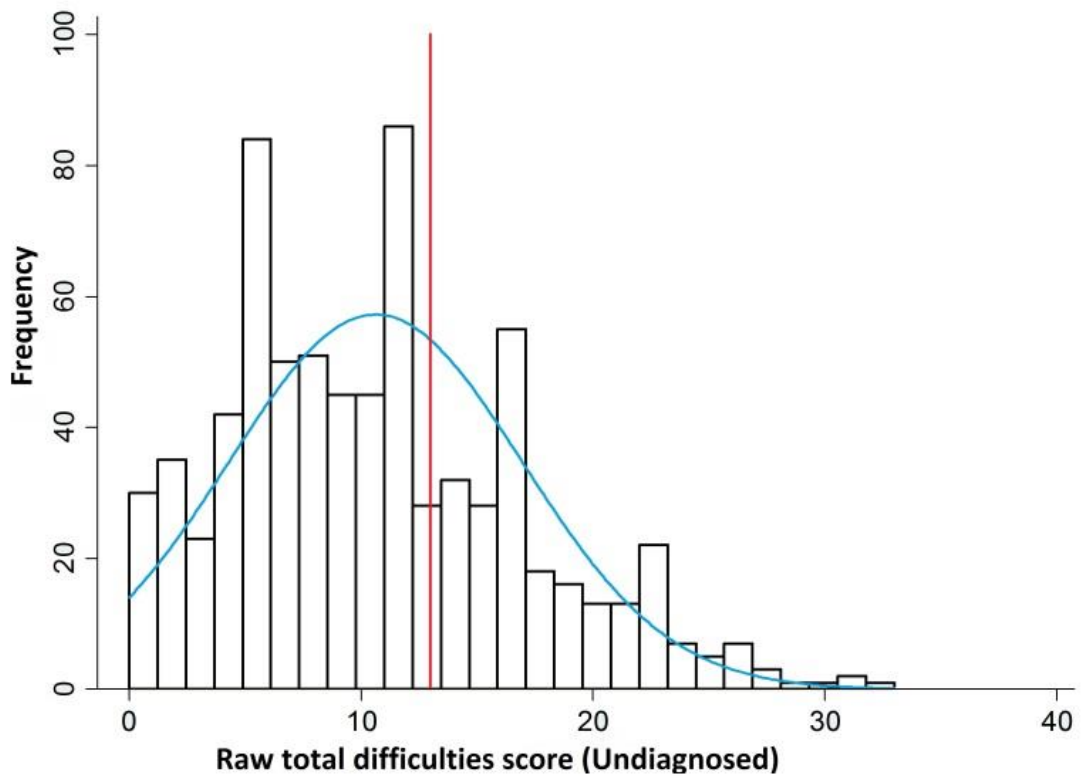

Figure A3.10 Histogram of the raw total difficulties scores at age 7 years for the children of undiagnosed mothers
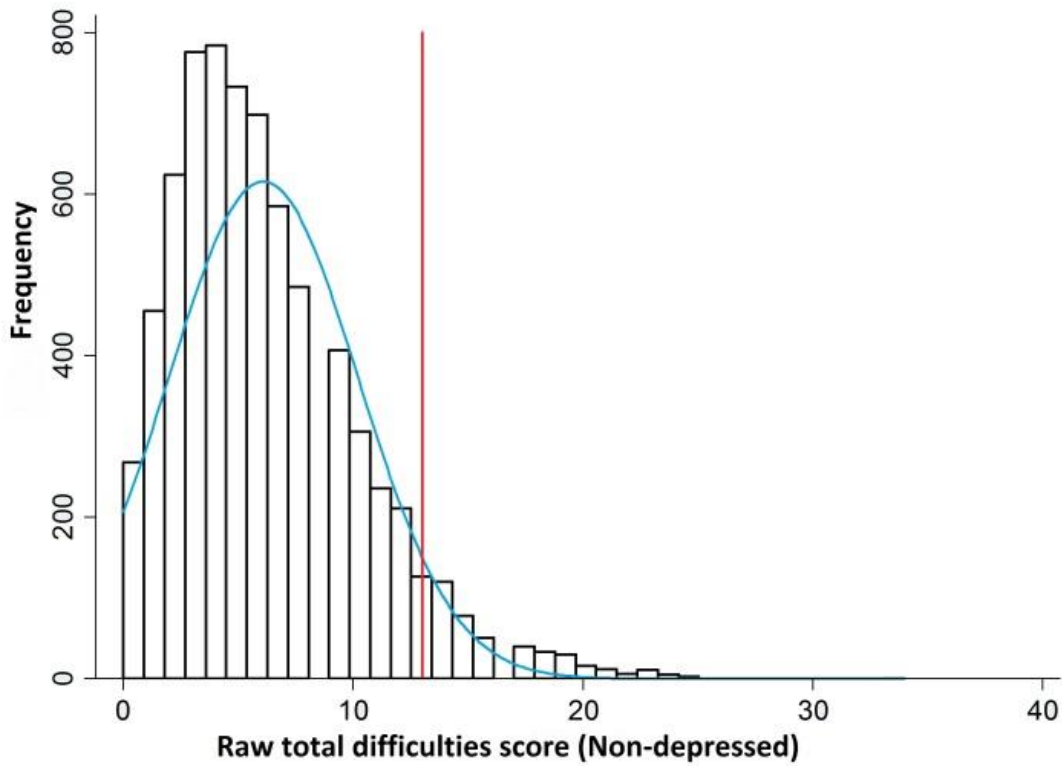
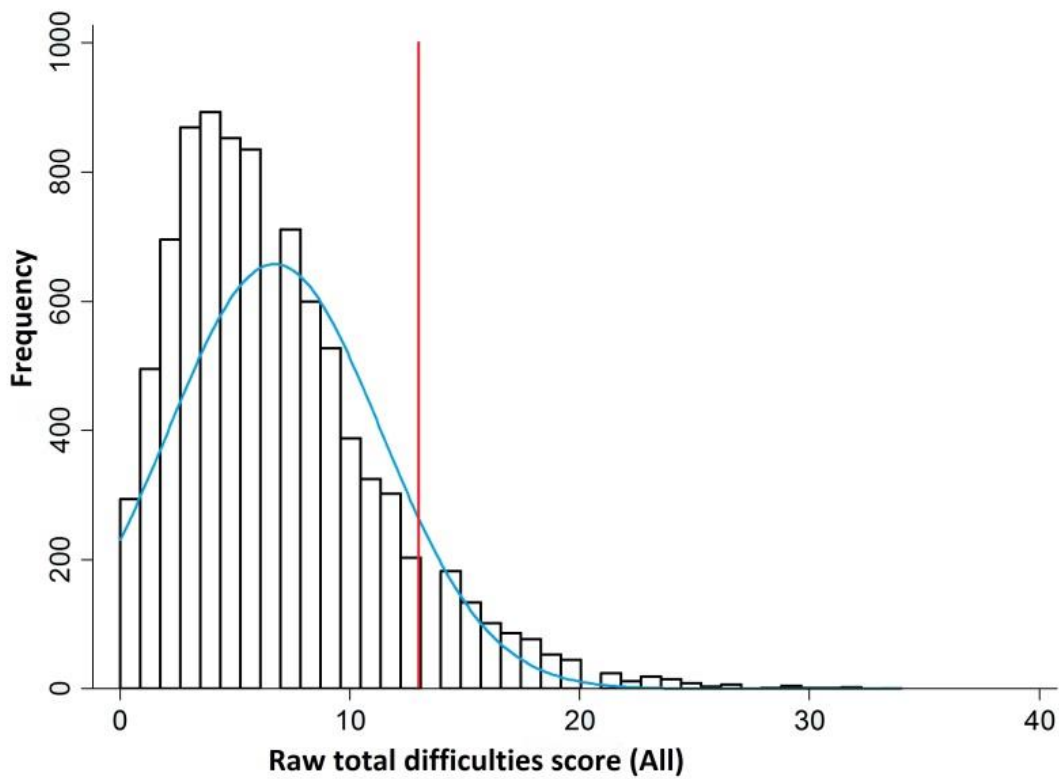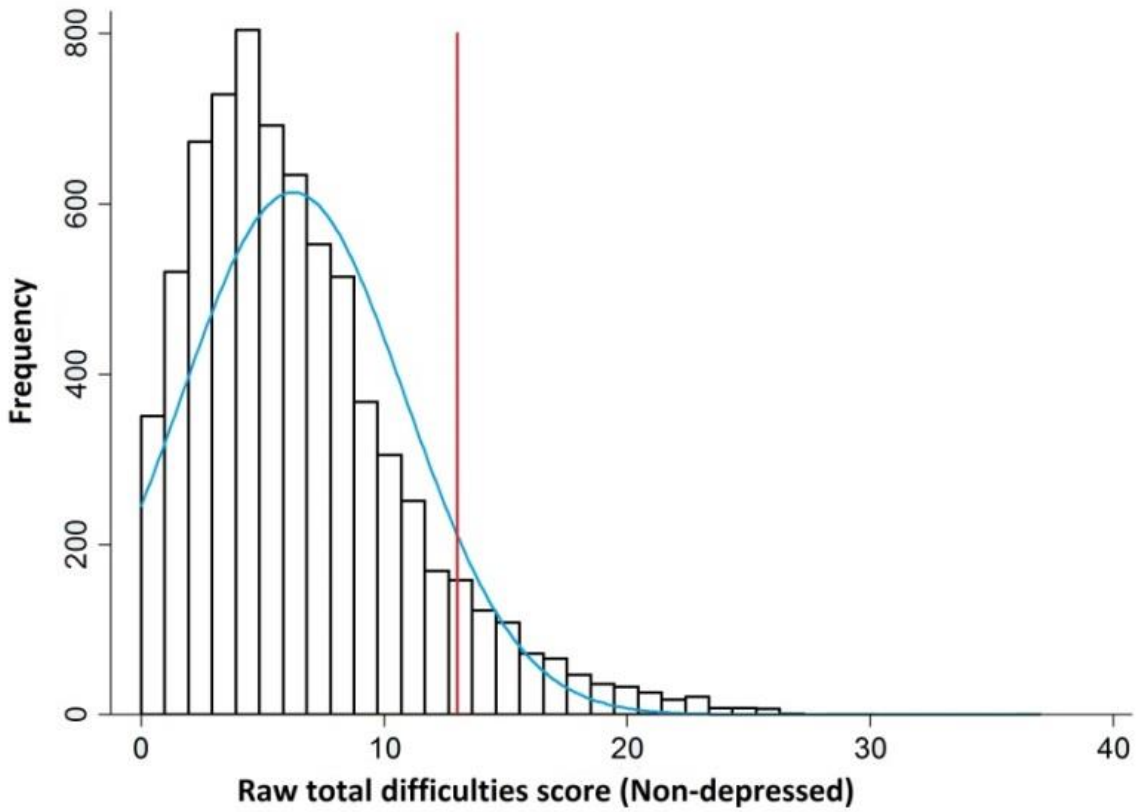Figure A3.11 Histogram of the raw total difficulties scores at age 7 years for children of non-depressed mothers
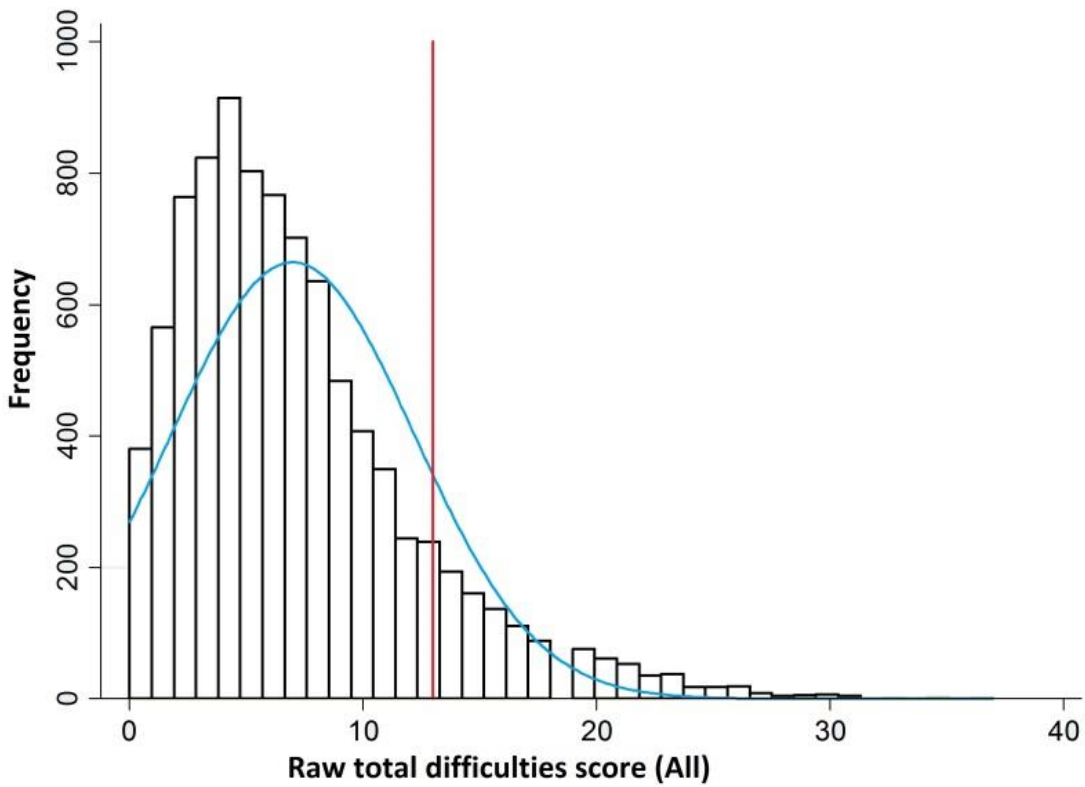
Figure A3.12 Histogram of the raw total difficulties scores at age 7 years for all children

Figure A3.13 shows an example of the diagnostic plots generated for each of the imputed data sets. These were generated for each of the models relating to externalising behaviours, internalising behaviours and in relation to the total difficulties score. An example for each of the models is included (externalising: Figure A3.13, internalising: Figure A3.14, total difficulties score: Figure A3.15).
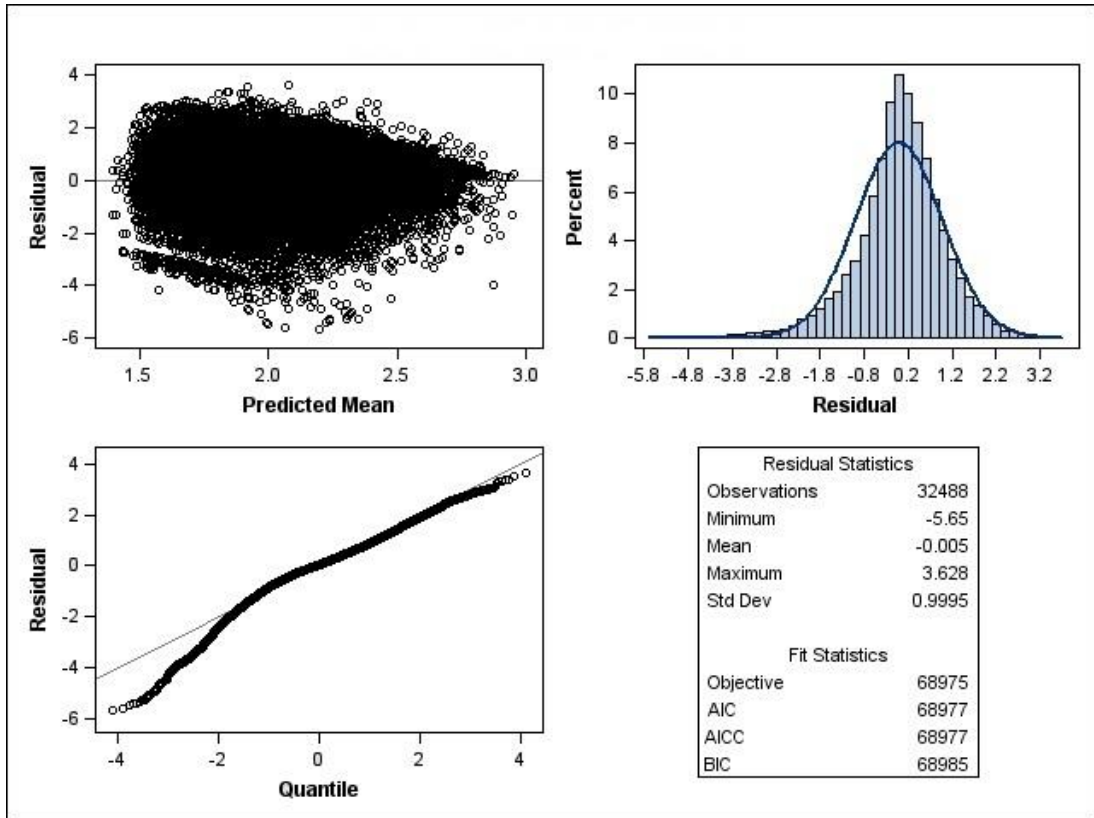
Figure A3.13 Diagnostic plots for imputed dataset 1: externalising behaviours model
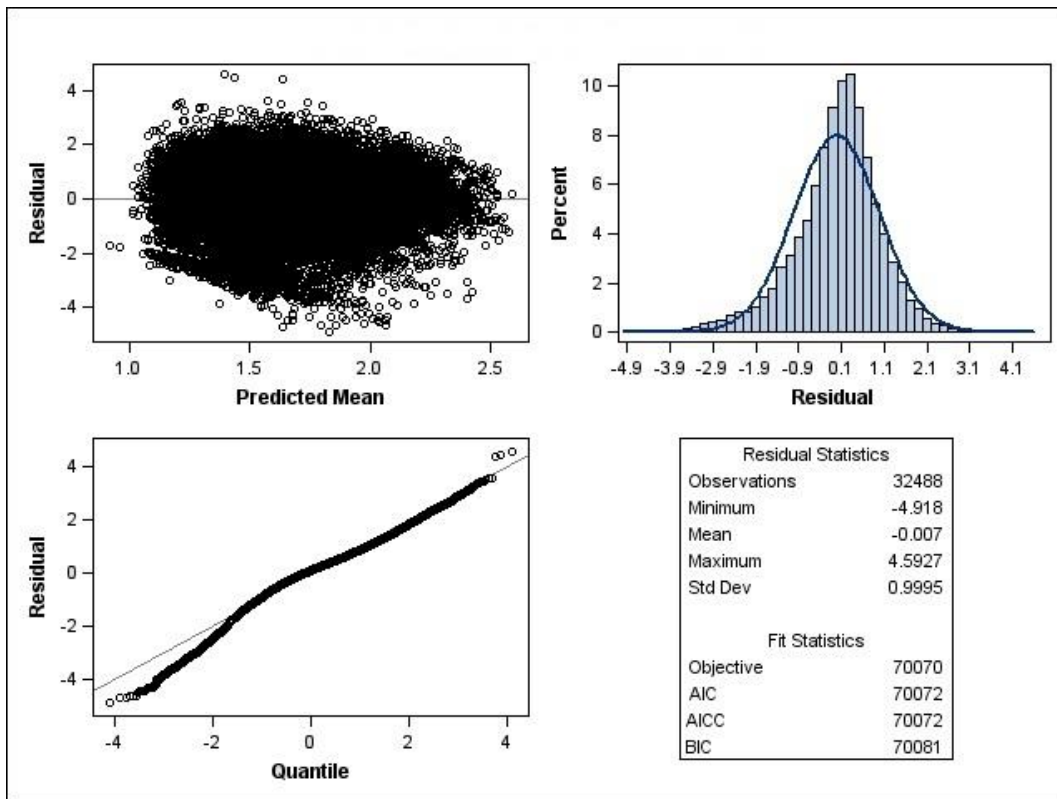
Figure A3.14 Diagnostic plots for imputed dataset 1: internalising behaviours model
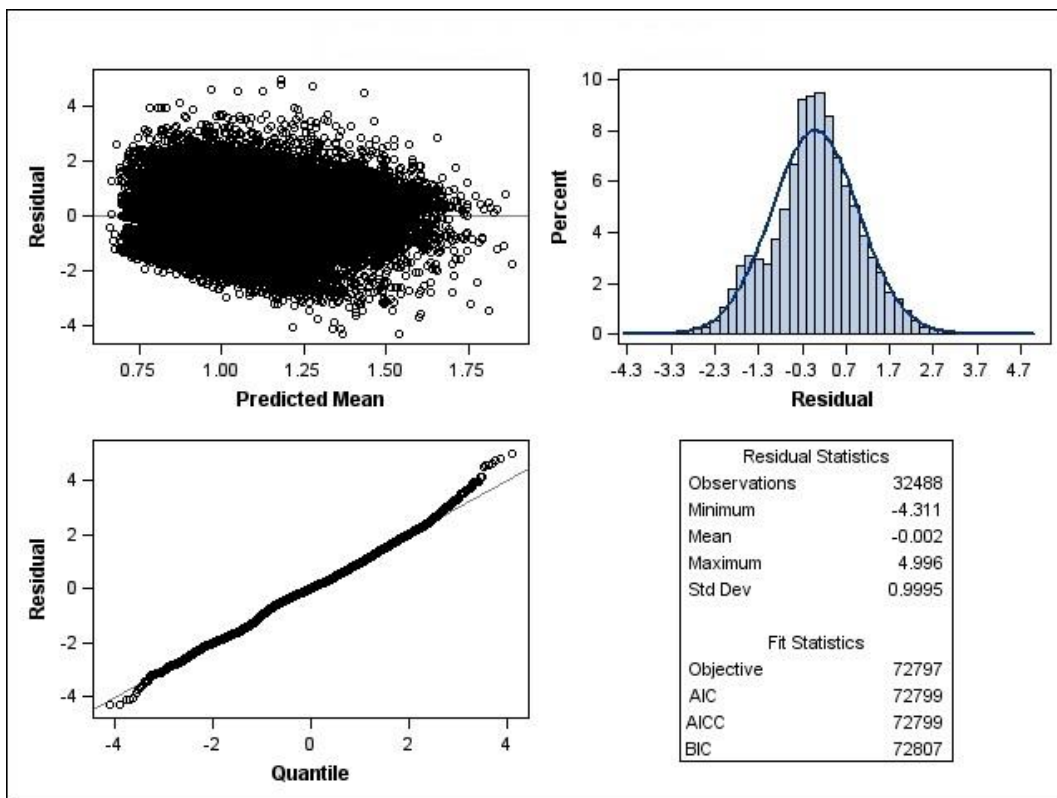
Figure A3.15 Diagnostic plots for imputed dataset 1: total difficulties model

Diagnostic plots were also generated for each of the final models.
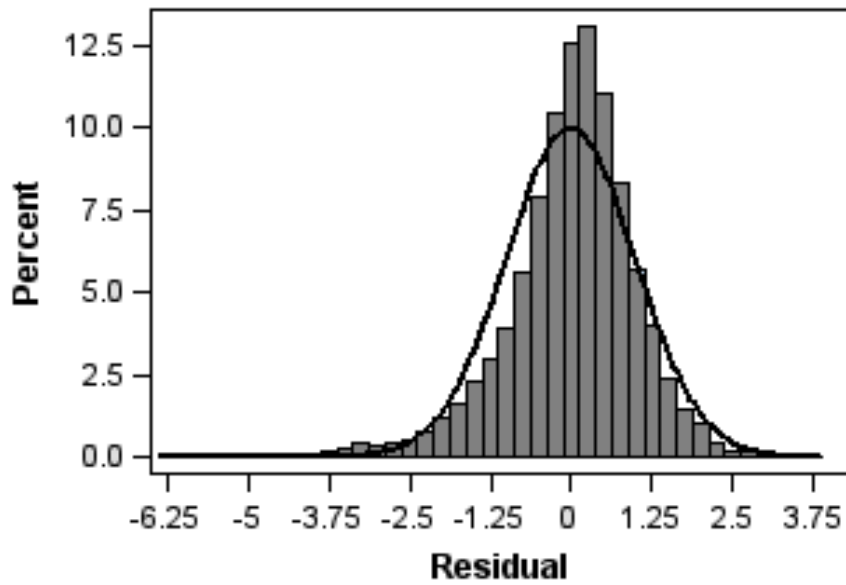
Figure A3.16 A histogram of the residuals versus the fitted values for the total difficulties model
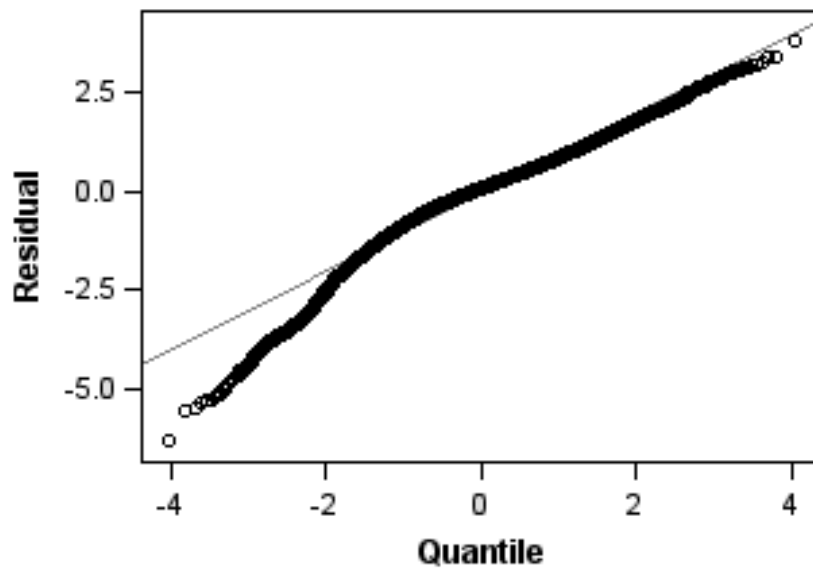
Figure A3.17 A Q-Q plot of the residuals versus the expected order statistics of the standard normal distribution for the total difficulties model
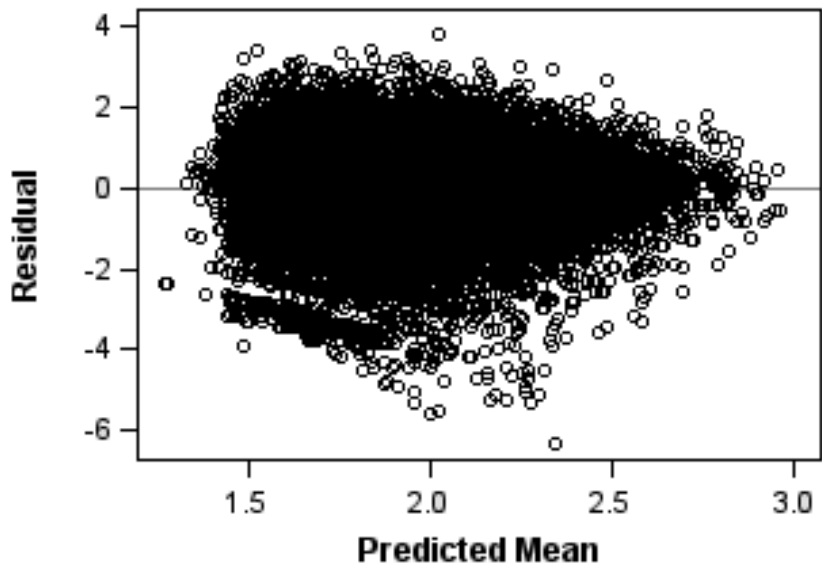
Figure A3.18 A scatter plot of the residuals versus the fitted values for the total difficulties model
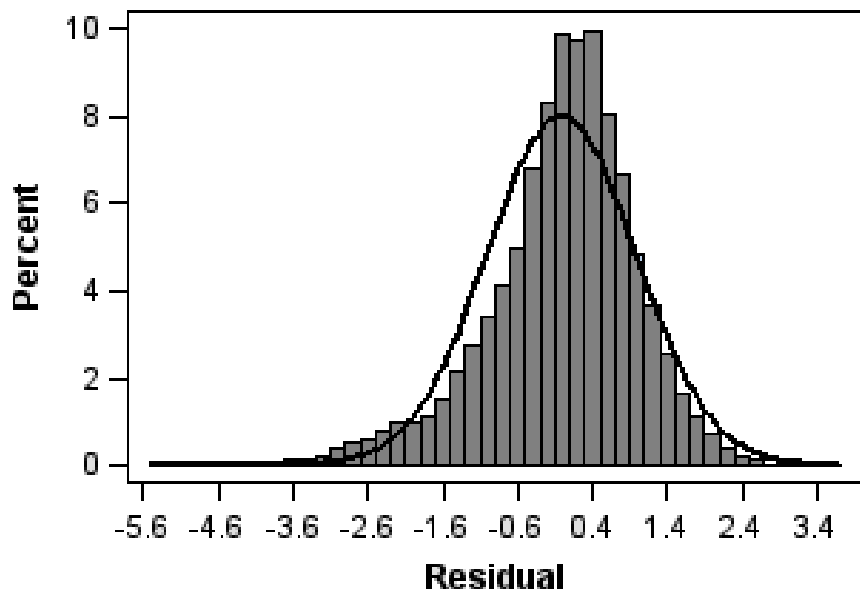
Figure A3.19 A histogram of the residuals versus the fitted values for the externalising difficulties model
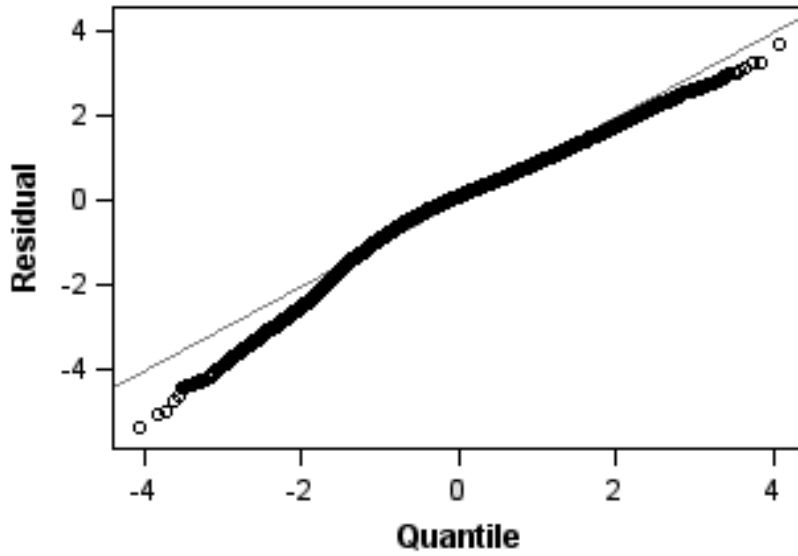
Figure A3.20 A Q-Q plot of the residuals versus the expected order statistics of the standard normal distribution for the externalising difficulties model
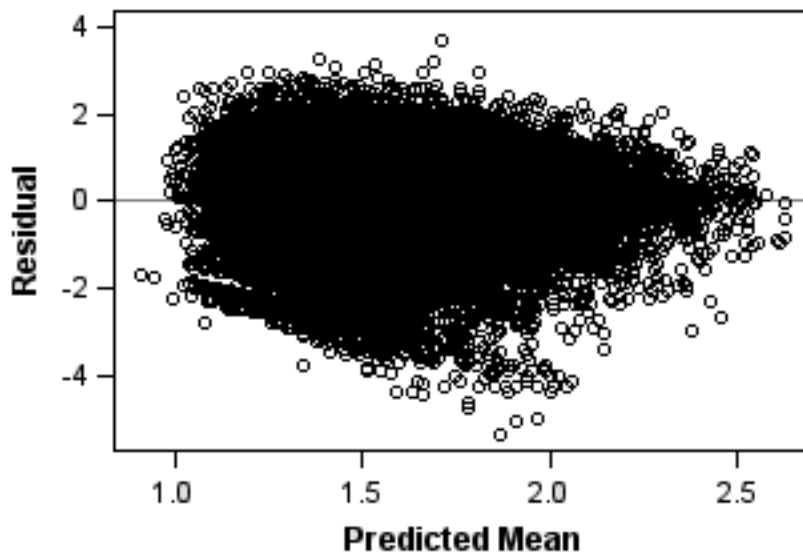
Figure A3.21 A scatter plot of the residuals versus the fitted values for the externalising difficulties model
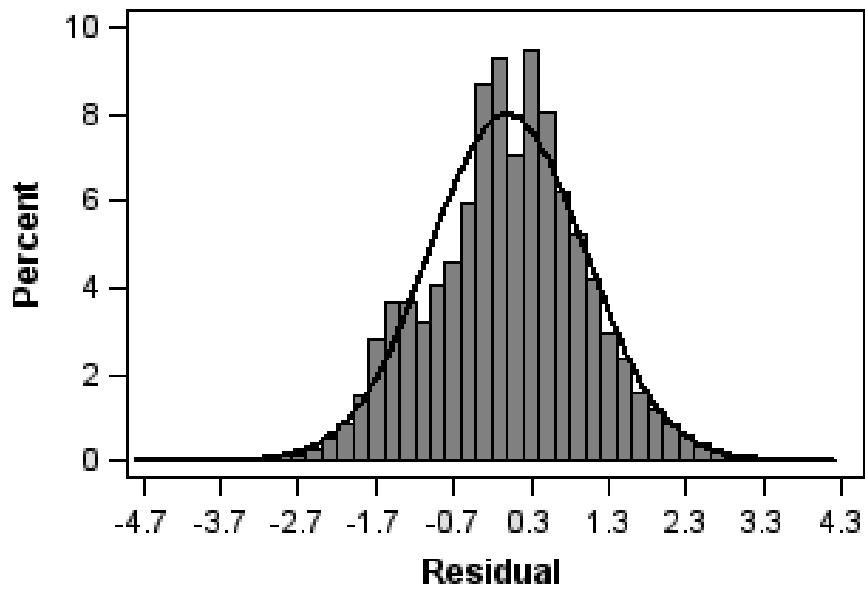
Figure A3.22 A histogram of the residuals versus the fitted values for the internalising difficulties model
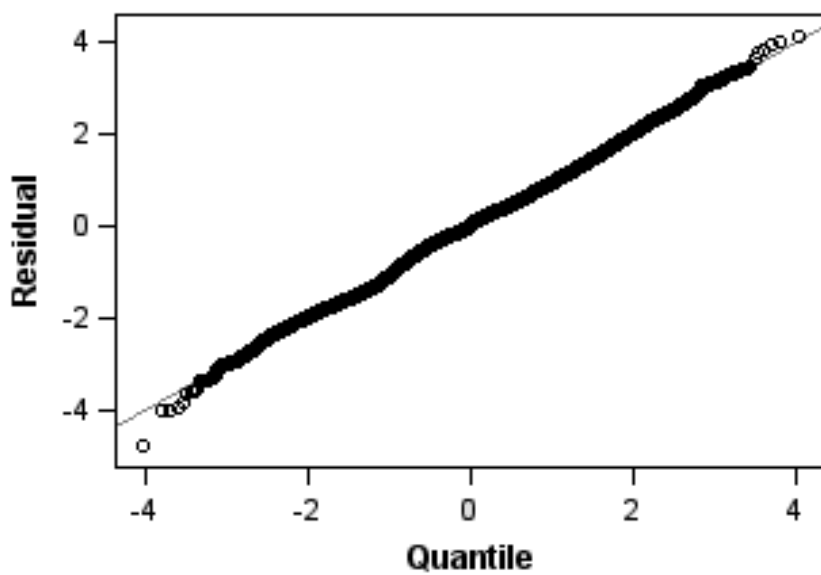
Figure A3.23 A Q-Q plot of the residuals versus the expected order statistics of the standard normal distribution for the internalising difficulties model
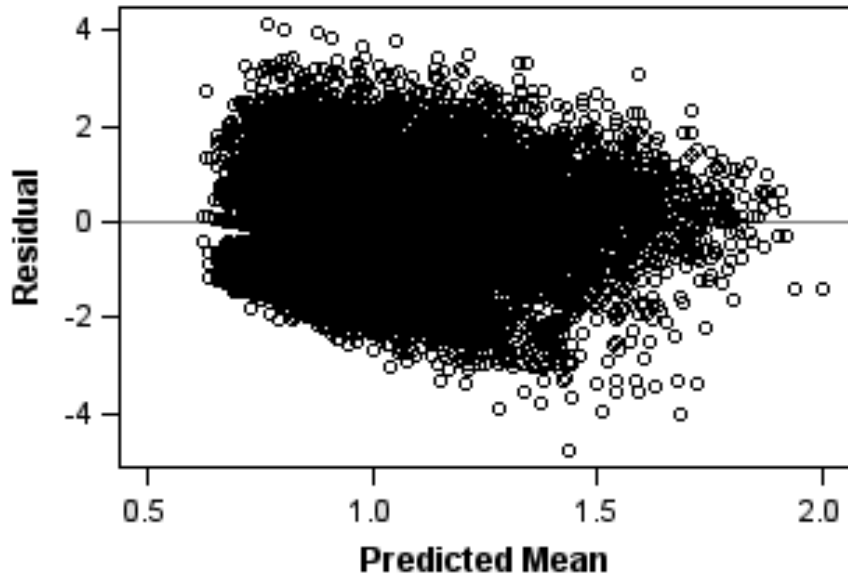
Figure A3.24 A scatter plot of the residuals versus the fitted values for the externalising difficulties model

Appendix 4: Chapter 6 A scoping review

Example search strategy targeting early behavioural problems and educational attainment measured by UK standard educational certificates:

1. exp child$/
2. Child$.ab,kw,ti.
3. early years.ab,kw,ti.
4. infant.ab,kw,ti.
5. toddler.ab,kw,ti.
6. exp development/
7. development.ab,kw,ti.
8. behavio$.ab,kw,ti.
9. conduct.ab,kw,ti.
10. hyperactiv$.ab,kw,ti.
11. peer problems.ab,kw,ti.
12. peer.ab,kw,ti.
13. emotion$.ab,kw,ti.
14. attention.ab,kw,ti.
15. inattention.ab,kw,ti.
16. exp education$/
17. education$.ab,kw,ti.
18. attainment.ab,kw,ti.
19. academic.ab,kw,ti.
20. higher education.ab,kw,ti.
21. further education.ab,kw,ti.
22. high school.ab,kw,ti.
23. college.ab,kw,ti.
24. GCSE$.ab,kw,ti.
25. A level.ab,kw,ti.
26. 1 or 2 or 3 or 4 or 5
27. 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15
28. 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25
29. 26 and 27 and 28
30. problems.ab,kw,ti.
31. issue$.ab,kw,ti.
32. difficult$.ab,kw,ti.
33. 30 or 31 or 32
34. 29 and 33

Abbreviations and Definitions

ADHD	Attention Deficit hyperactivity Disorder
AIC	Akaike's Information Criteria
AMSTAR	Assessment of Multiple Systematic Reviews
APA	American Psychological Association
BAS	British Ability Scales
BCS	British Cohort Study
BDI	Beck Depression Inventory
CAAS	Children's Attention and Adjustment Survey
CBCL	Child Behaviour Check List
CBT	Cognitive Behavioural Therapy
CHD	Christchurch Health and Development Study
CRO	Criminal Records Office
DAWBA	Development and Well-being Assessment
DSM	Diagnostic and Statistical Manual of Mental Disorders
DWP	Department of Work and Pensions (UK)
EPDS	Edinburgh Postnatal Depression Scale
EU	European Union
GCSE	General Certificate of Education (UK)
HAMD	Hamilton Rating Scale for Depression
HTA	Health Technology Assessment
MCMC	Markov Chain Monte Carlo
MCS	Millennium Cohort Study
MeSH	Medical Subject Headings
MI	Multiple Imputation
NICE	National Institute for Clinical Excellence
NIHR	National Institute for Health Research
NLSY	National Longitudinal Survey of Youth
NVQ	National Vocational Qualification
OECD	Organisation for Economic Co-operation and Development

OQAQ	Overview Quality Assessment Questionnaire
OR	Odds Ratio
PHQ-9	Patient Health Questionnaire-9
PND	Postnatal depression
QALY(s)	Quality Adjusted Life Year(s)
RCM	Royal College of Midwives
RCT	Randomised Controlled Trial
RECAP	Retrospective Child and Adolescent Psychopathology
SDQ	Strengths and Difficulties Questionnaire
TRF	Teacher Report Form
UK	United Kingdom
US(A)	United States (of America)
WF-R	Woodcock Johnson-Revised

Diagnosed	Mothers reporting that they had been diagnosed with depression by a doctor and were receiving treatment at the time of the questionnaire
Undiagnosed	Mothers reporting that they were not receiving treatment for depression but who identified as depressed on a self-reported scale
Non-depressed	Mothers reporting that they were not receiving treatment for depression and who did not identify as depressed on a self-reported scale

References

4CHILDREN 2011. Suffering in Silence. London: 4Children.

ACHENBACH, T. M. 1978. The Child Behavior Profile: I. Boys aged 6--11. *J Consult Clin Psychol,* 46**,** 478-88.

ACHENBACH, T. M., HOWELL, C. T., QUAY, H. C. & CONNERS, C. K. 1991. National survey of problems and competencies among four- to sixteen-year-olds: parents' reports for normative and clinical samples. *Monogr Soc Res Child Dev,* 56**,** 1-131.

Achenbach, T.M., & Rescorla, L.A. (2000). *Manual for the ASEBA Preschool forms and profiles.* Burlington, VT: University of Vermont Department of Psychiatry.

ADAMS, J. W., SNOWLING, M. J., HENNESSY, S. M. & KIND, P. 1999. Problems of behaviour, reading and arithmetic: assessments of comorbidity using the Strengths and Difficulties Questionnaire. *Br J Educ Psychol,* 69 (Pt 4)**,** 571-85.

AGNAFORS, S., COMASCO, E., BLADH, M., SYDSJO, G., DEKEYSER, L., ORELAND, L. & SVEDIN, C. G. 2013. Effect of gene, environment and maternal depressive symptoms on pre-adolescence behavior problems - a longitudinal study. *Child Adolesc Psychiatry Ment Health,* 7**,** 10.

AGNAFORS, S., SYDSJO, G., DEKEYSER, L. & SVEDIN, C. G. 2013. Symptoms of depression postpartum and 12 years later-associations to child mental health at 12 years of age. *Matern Child Health J,* 17**,** 405-14.

AINSWORTH, M. D. 1969. Object relations, dependency, and attachment: a theoretical review of the infant-mother relationship. *Child Dev,* 40**,** 969-1025.

AINSWORTH, M.S., BLEHAR, M.C., WATERS, E., & WALL, S. 1978. *Patterns of attachment: A psychological study of the strange situation*. Oxford, England: Lawrence Erlbaum.

ALATUPA, S., PULKKI-RABACK, L., HINTSANEN, M., ELOVAINIO, M., MULLOLA, S. & KELTIKANGAS-JARVINEN, L. 2013. Disruptive behavior in childhood and socioeconomic position in adulthood: a prospective study over 27 years. *International Journal of Public Health,* 58**,** 247-256.

ALDER, E. 2007. *NAPD: the national audit of the detection and management of postnatal depression*. NHS Scotland.

ALI, S. 2012. Relationships among Perceived Teachers' and Parental Behavior and Adolescent Outcomes: A Cross-Cultural Literature Review. *European Psychiatry,* 27.

ALLISON, P.D. 2002. *Missing data. Quantitative Applications in the Social Sciences*. Thousand Oaks, CA: Sage Publications

ALMOND, P. 2009. Postnatal depression : A global public health perspective. *Perspectives in Public Health,* 129**,** 221-227.

ALTMAN, D. G. & BLAND, J. M. 2007. Missing data. *BMJ,* 334**,** 424.

AMERICAN PSYCHIATRIC ASSOCIATION. 1994. *Diagnostic and statistical manual of mental disorders 4th ed*. Washington, DC: Author

AMERICAN PSYCHIATRIC ASSOCIATION. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

ANDERSON, L. M., SHINN, C., FULLILOVE, M. T., SCRIMSHAW, S. C., FIELDING, J. E., NORMAND, J., CARANDE-KULIS, V. G. & TASK FORCE ON COMMUNITY PREVENTIVE, S. 2003. The effectiveness of early childhood development programs. A systematic review. *Am J Prev Med,* 24**,** 32-46.

ANTASI, A., & URBINA, S. (1997). Psychological testing (7th ed.). Upper Saddle River, NJ: Prentice Hall

APPLEBY, L., WARNER, R., WHITTON, A. & FARAGHER, B. 1997. A controlled study of fluoxetine and cognitive-behavioural counselling in the treatment of postnatal depression. *BMJ,* 314**,** 932-6.

ARKSEY, H. & O'MALLEY, L. 2005. Scoping studies: towards a methodological framework. *International Journal of Social Research Methodology,* 8**,** 19-32.

ARMIJO-OLIVO, S., STILES, C. R., HAGEN, N. A., BIONDO, P. D. & CUMMINGS, G. G. 2012. Assessment of study quality for systematic reviews: a comparison of the Cochrane Collaboration Risk of Bias Tool and the Effective Public Health Practice Project Quality Assessment Tool: methodological research. *J Eval Clin Pract,* 18**,** 12-8.

ARROLL, B., GOODYEAR-SMITH, F., KERSE, N., FISHMAN, T. & GUNN, J. 2005. Effect of the addition of a "help" question to two screening questions on specificity for diagnosis of depression in general practice: diagnostic validity study. *BMJ,* 331**,** 884.

ASHMAN, S. B., DAWSON, G. & PANAGIOTIDES, H. 2008. Trajectories of maternal depression over 7 years: relations with child psychophysiology and behavior and role of contextual risks. *Dev Psychopathol,* 20**,** 55-77.

ATKINSON, T., BOURGUIGNON, F., O'DONOGHUE, C., SUTHERLAND, H. & UTILI, F. 2002. Microsimulation of social policy in the European Union: Case study of a European Minimum Pension. *Economica,* 69**,** 229-243.

BABINSKI, L. M., HARTSOUGH, C. S. & LAMBERT, N. M. 1999. Childhood conduct problems, hyperactivity-impulsivity, and inattention as predictors of adult criminal activity. *J Child Psychol Psychiatry,* 40**,** 347-55.

BARKER, D.J.P. 1998. *Mothers, Babies and Health in Later Life, (2nd Edition).* Edinburgh: Churchill Livingstone.

BAKER, D., NORTH, K. & TEAM, A. S. 1999. Does employment improve the health of lone mothers? *Social Science & Medicine,* 49**,** 121-131.

BARKER, D. J. 2001. A new model for the origins of chronic disease. *Med Health Care Philos,* 4**,** 31-5.

BARKLEY, R. A., FISCHER, M., SMALLISH, L. & FLETCHER, K. 2006. Young adult outcome of hyperactive children: adaptive functioning in major life activities. *J Am Acad Child Adolesc Psychiatry,* 45**,** 192-202.

BARTLEY, M. 2004. *Health Inequality: An introduction to theories, concepts and methods,* Oxford, Polity Press.

BARNETT, S. 1995. Long Term Effects of Early Childhood Programs on Cognitive and School Outcomes. *The Future of Children,* 5, 25–50.

BARTLEY, M., KELLY, Y., SCHOON, I. & HOPE, S. 2004. Parent health. In S. Dex & H. Joshi (eds). *Millennium Cohort Study First Survey: A User's Guide to Initial Findings*. Centre for Longitudinal Studies, Institute of Education, University of London, London, UK. pp. 140–158

BATTY, G. D., DEARY, I. J., SCHOON, I. & GALE, C. R. 2007. Mental ability across childhood in relation to risk factors for premature mortality in adult life: the 1970 British Cohort Study. *J Epidemiol Community Health,* 61**,** 997-1003.

BECK, A. T., STEER, R. A., BALL, R. & RANIERI, W. 1996. Comparison of Beck Depression Inventories -IA and -II in psychiatric outpatients. *J Pers Assess,* 67**,** 588-97.

BECK, C. T. & GABLE, R. K. 2000. Postpartum Depression Screening Scale: development and psychometric testing. *Nurs Res,* 49**,** 272-82.

BEDFORD, A. & FOULDS, G. A. 1978. *Manual of rhe Delusions-Symptoms-States Inventory (DSSI).* Windsor: NFER-Nelson.

BELMAKER, R. H. & AGAM, G. 2008. Major depressive disorder. *N Engl J Med,* 358**,** 55-68.

BELSKY, J., MELHUISH, E., BARNES, J., LEYLAND, A. H. & ROMANIUK, H. 2006. Effects of Sure Start local programmes on children and families: early findings from a quasi-experimental, cross sectional study. *BMJ,* 332**,** 1476.

BENGSTON, V. L. & ALLEN, K. R. 1993. The life course perspective applied to families over time. *In:* BOSS, P., DOHERTY, W., LAROSSA, R., SCHUMM, W. & STEINMETZ, S. (eds.) *Sourcebook of family theories and methods: A contextual approach* New York: Plenum Press.

BEN-SHLOMO, Y. & KUH, D. 2002. A life course approach to chronic disease epidemiology: conceptual models, empirical challenged and interdisciplinary perspectives. *International Journal of Epidemiology,* 31**,** 285-293.

BEN-SHLOMO, Y. & KUH, D. 2002. A life course approach to chronic disease epidemiology: conceptual models, empirical challenges and interdisciplinary perspectives. *Int J Epidemiol,* 31**,** 285-93.

BHATTACHARYA, A. 2010. Children and Adolescents From Poverty and Reading Development: A Research Review. *Reading & Writing Quarterly,* 26**,** 115-139.

BLACK, M. B., & MATULA, K. (1999).Essentials of Bayley Scales of Infant Development II Assessment. New York: John Wiley & Sons, Inc.

BLAND, M. 2000. *An Introduction to Medical Statistics, Third Edition.* Oxford: Oxford University Press

BLANE, D. 1999. The lifecourse, the social gradient and health. *In:* MARMOT, M. & WILKINSON, R. G. (eds.) *Social Determinants of Health.* Oxford: Oxford University Press.

BLANE, D., DAVEY SMITH, G. & BARTLEY, M. 1993. Social selection: what does it contribute to social-class differences in health? *Sociology of Health and Illness,* 15**,** 1-15.

BLIER, P., KELLER, M. B., POLLACK, M. H., THASE, M. E., ZAJECKA, J. M. & DUNNER, D. L. 2007. Preventing recurrent depression: long-term treatment for major depressive disorder. *J Clin Psychiatry,* 68**,** e06.

BOATH, E. H., PRYCE, A. J. & COX, J. L. 1998. Postnatal depression: The impact on the family. *Journal of Reproductive and Infant Psychology* 16**,** 199-203.

BODNER, T. E. 2008. What Improves with Increased Missing Data Imputations? *Structural Equation Modeling-a Multidisciplinary Journal,* 15**,** 651-675.

BOND, R. & SAUNDERS, P. 1999. Routes of success: influences on the occupational attainment of young British males. *Br J Sociol,* 50**,** 217-49.

BOR, W., MCGEE, T. R. & FAGAN, A. A. 2004. Early risk factors for adolescent antisocial behaviour: an Australian longitudinal study. *Aust N Z J Psychiatry,* 38**,** 365-72.

BOWLBY, J. 1969. *Attachment. Attachment and Loss: Vol. 1. Loss*. New York: Basic Books

BOYLE, M. H. & PICKLES, A. 1997. Maternal depressive symptoms and ratings of emotional disorder symptoms in children and adolescents. *J Child Psychol Psychiatry,* 38**,** 981-92.

BOYLE, M. H. & PICKLES, A. R. 1997. Influence of maternal depressive symptoms on ratings of childhood behavior. *J Abnorm Child Psychol,* 25**,** 399-412.

BRACKEN, B. A. (1984). *Bracken basic concept scale*. San Antonio: Harcourt Assessments.

BRACKEN, B. A. (2002). *Bracken school readiness assessment*. San Antonio, TX: The Psychological Corporation.

BRAVEMAN, P. & BARCLAY, C. 2009. Health disparities beginning in childhood: a life-course perspective. *Pediatrics,* 124 Suppl 3**,** S163-75.

BREINHOLST, S., ESBJORN, B. H., REINHOLDT-DUNNE, M. L. & STALLARD, P. 2012. CBT for the treatment of child anxiety disorders: A review of why parental involvement has not enhanced outcomes. *Journal of Anxiety Disorders,* 26**,** 416-424.

BRENNAN, P. A., HAMMEN, C., ANDERSEN, M. J., BOR, W., NAJMAN, J. M. & WILLIAMS, G. M. 2000. Chronicity, severity, and timing of maternal depressive symptoms: relationships with child outcomes at age 5. *Dev Psychol,* 36**,** 759-66.

BRESLAU, J., MILLER, E., BRESLAU, N., BOHNERT, K., LUCIA, V. & SCHWEITZER, J. 2009. The impact of early behavior disturbances on academic achievement in high school. *Pediatrics,* 123**,** 1472-6.

BRESLAU, N., CHILCOAT, H., DELDOTTO, J., ANDRESKI, P. & BROWN, G. 1996. Low birth weight and neurocognitive status at six years of age. *Biol Psychiatry,* 40**,** 389-97.

BRETHERTON, I. 1985. Attachment theory. Retrospect and prospect. In I. Bretherton & E. Waters (Eds.), *Growing points of attachment theory and research, Monographs of the Society for Research in Child Development*, 50(1-2, Serial No. 209), 3-3

BRETHERTON, I. 1992. The origins of attachment theory: John Bowlby and Mary Ainsworth. *Developmental Psychology, 28,* 759–775.

BRETHERTON, I., & MUNHOLLAND, K. A. 2008. Internal working models in attachment relationships: Elaborating a central construct in attachment theory. In J. CASSIDY & P. R. SHAVER (Eds). *Handbook of attachment: Theory, research, and clinical applications (2nd ed.).* New York, NY: Guilford Press pp. 102–127

BRIGGS, A., CLAXTON, K. & SCULPHER, M. 2006. *Decision modelling for health economic evaluation.,* Oxford, Oxford University Press.

BROOKS-GUNN, J., BRITTO, L. P. R. & BRADY, C. 1999. Struggling to make ends meet: Poverty and child development. *In:* LAMB, M. E. (ed.) *Parenting and Child Development in "Nontraditional" Families.* Mahwah, NJ: Lawrence Erlbaum Associates.

BROOKS-GUNN, J., HAN, W. J. & WALDFOGEL, J. 2002. Maternal employment and child cognitive outcomes in the first three years of life: the NICHD Study of Early Child Care. National Institute of Child Health and Human Development. *Child Dev,* 73**,** 1052-72.

BROTH, M. R., GOODMAN, S. H., HALL, C. & RAYNOR, L. C. 2004. Depressed and well mothers' emotion interpretation accuracy and the quality of mother-infant interaction. *Infancy,* 6**,** 37-55.

BROWN, G. W. & BIFULCO, A. 1990. Motherhood, employment and the development of depression. A replication of a finding? *Br J Psychiatry,* 156**,** 169-79.

BROWN, G. W. & MORAN, P. M. 1997. Single mothers, poverty and depression. *Psychological Medicine,* 27**,** 21-33.

BUSS, A.H. & PLOMIN, R. (1984). Temperament: Early developing personality traits. Hillsdale, NJ: Lawrence Erlbaum.

BUXTON, M. J., DRUMMOND, M. F., VAN HOUT, B. A., PRINCE, R. L., SHELDON, T. A., SZUCS, T. & VRAY, M. 1997. Modelling in economic evaluation: an unavoidable fact of life. *Health Econ,* 6**,** 217-27.

BYNG-HALL, J. 2005. Attachment and human survival. *Journal of Family Therapy,* 27**,** 97-98.

BYNNER, J., FERRI, E. & SHEPHERD, P. 1997. *Twenty something in the 1990s: Getting on, getting by, getting nowhere.,* Aldershot, Ashgate.

CAMPBELL, F., RAMSEY, C., PUNGELLO, E., SPARLING, J. & MILLER-JOHNSON, S. 2002. Early Childhood Education: Young Adult Outcomes From the Abecedarian Project. . *Applied Developmental Science,* 6**,** 42-57.

CAMPBELL, S. B. 1995. Behavior problems in preschool children: a review of recent research. *J Child Psychol Psychiatry,* 36**,** 113-49.

CAMPBELL, S. B. & EWING, L. J. 1990. Follow-up of hard-to-manage preschoolers: adjustment at age 9 and predictors of continuing symptoms. *J Child Psychol Psychiatry,* 31**,** 871-89.

CAMPBELL, S. B., EWING, L. J., BREAUX, A. M. & SZUMOWSKI, E. K. 1986. Parent-referred problem three-year-olds: follow-up at school entry. *J Child Psychol Psychiatry,* 27**,** 473-88.

CAMPBELL, S. B., SHAW, D. S. & GILLIOM, M. 2000. Early externalizing behavior problems: toddlers and preschoolers at risk for later maladjustment. *Dev Psychopathol,* 12**,** 467-88.

CARLSON, E. A. 1998. A prospective longitudinal study of attachment disorganization/disorientation. *Child Dev,* 69**,** 1107-28.

CARLSON, E. A., SROUFE, L. A. & EGELAND, B. 2004. The construction of experience: a longitudinal study of representation and behavior. *Child Dev,* 75**,** 66-83.

CARNEIRO, P. & HECKMAN, J. J. 2003. Human Capital Policy. *In:* HECKMAN, J. J., KRUEGER, A. B. & FRIEDMAN, B. (eds.) *Inequality in America: What Role for Human Capital Policies?* . Cambridge, MA: MIT Press.

CASPI, A., ENTNER WRIGHT, B. R., MOFFITT, T. E. & SILVA, P. A. 1998. Early failure in the labour market: Childhood and adolescent predictors of unemployment in the transition to adulthood. *American Sociological Review,* 63**,** 424-451.

CHABROL, H., TEISSEDRE, F., SAINT-JEAN, M., TEISSEYRE, N., ROGE, B. & MULLET, E. 2002. Prevention and treatment of post-partum depression: a controlled randomized study on women at risk. *Psychol Med,* 32**,** 1039-47.

CHAIMAY, B., THINKHAMROP, B. & THINKHAMROP, J. 2006. Risk factors associated with language development problems in childhood--a literature review. *J Med Assoc Thai,* 89**,** 1080-6.

CHAN, S. & LEVY, V. 2004. Postnatal depression: a qualitative study of the experiences of a group of Hong Kong Chinese women. *J Clin Nurs,* 13**,** 120-3.

CHAN, S. W., LEVY, V., CHUNG, T. K. & LEE, D. 2002. A qualitative study of the experiences of a group of Hong Kong Chinese women diagnosed with postnatal depression. *J Adv Nurs,* 39**,** 571-9.

CHEN, E., MATTHEWS, K. A. & BOYCE, W. T. 2002. Socioeconomic differences in children's health: how and why do these relationships change with age? *Psychol Bull,* 128**,** 295-329.

CHEN, H. Y. & ZHU, J. J. 2008. Factor invariance between genders of the Wechsler Intelligence Scale for Children - Fourth edition. *Personality and Individual Differences,* 45**,** 260-266.

CHEN, W. J., FARAONE, S. V., BIEDERMAN, J. & TSUANG, M. T. 1994. Diagnostic accuracy of the Child Behavior Checklist scales for attention-deficit hyperactivity

disorder: a receiver-operating characteristic analysis. *J Consult Clin Psychol, 62,* 1017-25.

CHISHOLM, D. & HEALEY, A.T. & KNAPP, M.R.J. 1997. QALYs and mental health care. *Social Psychiatry and Psychiatric Epidemiology, 32,* 68-75

CICCHETTI, D., ROGOSCH, F. A. & TOTH, S. L. 2006. Fostering secure attachment in infants in maltreating families through preventive interventions. *Dev Psychopathol,* 18, 623-49.

CLARK, R., TLUCZEK, A. & WENZEL, A. 2003. Psychotherapy for postpartum depression: a preliminary report. *Am J Orthopsychiatry, 73,* 441-54.

COLEY, R. L. 1998. Children's socialization experiences and functioning in single-mother households: the importance of fathers and other men. *Child Dev, 69,* 219-30.

COLEY, R. L. & CHASE-LANSDALE, P. L. 1998. Adolescent pregnancy and parenthood. Recent evidence and future directions. *Am Psychol, 53,* 152-66.

CONGER, R. D., CONGER, K. J., ELDER, G. H., JR., LORENZ, F. O., SIMONS, R. L. & WHITBECK, L. B. 1992. A family process model of economic hardship and adjustment of early adolescent boys. *Child Dev, 63,* 526-41.

CONNELLY, R. 2013. MCS data note 2013/1: Interpreting test scores. [Online] [Accessed 12 October 2013] Available from http://www.cls.ioe.ac.uk/

CONRAD, M. & HAMMEN, C. 1989. Role of maternal depression in perceptions of child maladjustment. *J Consult Clin Psychol, 57,* 663-7.

CORNISH, A. M., MCMAHON, C. A., UNGERER, J. A., BARNETT, B., KOWALENKO, N. & TENNANT, C. 2006. Maternal depression and the experience of parenting in the second postnatal year. *Journal of Reproductive and Infant Psychology, 24,* 121-132.

COX, J. L., HOLDEN, J. M. & SAGOVSKY, R. 1987. Detection of postnatal depression. Development of the 10-item Edinburgh Postnatal Depression Scale. *Br J Psychiatry,* 150**,** 782-6.

COX, J. L., MURRAY, D. & CHAPMAN, G. 1993. A Controlled-Study of the Onset, Duration and Prevalence of Postnatal Depression. *British Journal of Psychiatry,* 163**,** 27-31.

CROSIER, T., BUTTERWORTH, P. & RODGERS, B. 2007. Mental health problems among single and partnered mothers. The role of financial hardship and social support. *Soc Psychiatry Psychiatr Epidemiol,* 42**,** 6-13.

CUNHA, F. & HECKMAN, J. 2007. The technology of skill formation. *American Economic Review,* 97**,** 31-47.

CUNHA, F. & HECKMAN, J. J. 2008. Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation. *Journal of Human Resources,* 43**,** 738-782.

CURRIE, M. L. & RADEMACHER, R. 2004. The pediatrician's role in recognizing and intervening in postpartum depression. *Pediatr Clin North Am,* 51**,** 785-801, xi.

DAVEY SMITH, G., BEN-SHLOMO, Y. & LYNCH, J. W. 2002. Life course approaches to inequalities in coronary heart disease risk. . *In:* STANSFELD, S. & MARMOT, M. (eds.) *Stress and the heart.* London: BMJ Books.

DEAULT, L. C. 2010. A systematic review of parenting in relation to the development of comorbidities and functional impairments in children with attention-deficit/hyperactivity disorder (ADHD). *Child Psychiatry Hum Dev,* 41**,** 168-92.

DEEKS, J. J., DINNES, J., D'AMICO, R., SOWDEN, A. J., SAKAROVITCH, C., SONG, F., PETTICREW, M. & ALTMAN, D. G. 2003. Evaluating non-randomised intervention studies. *Health Technol Assess,* 7**,** iii-x, 1-173.

DENNIS, C. L. 2003. The effect of peer support on postpartum depression: a pilot randomized controlled trial. *Can J Psychiatry,* 48**,** 115-24.

DENNIS, C. L. 2005. Psychosocial and psychological interventions for prevention of postnatal depression: systematic review. *BMJ,* 331**,** 15.

DENNIS, C. L. & CHUNG-LEE, L. 2006. Postpartum depression help-seeking barriers and maternal treatment preferences: a qualitative systematic review. *Birth,* 33**,** 323-31.

DENNIS, C. L. & CREEDY, D. 2004. Psychosocial and psychological interventions for preventing postpartum depression. *Cochrane Database Syst Rev***,** CD001134.

DENNIS, C. L. & DOWSWELL, T. 2013. Psychosocial and psychological interventions for preventing postpartum depression. *Cochrane Database Syst Rev,* 2**,** CD001134.

DENNIS, C. L. & HODNETT, E. 2007. Psychosocial and psychological interventions for treating postpartum depression. *Cochrane Database Syst Rev***,** CD006116.

Depression, teen mums and me. BBC 3. 19 March 2013. 2100 hrs.

DEPARTMENT FOR EDUCATION AND SKILLS. 2007. Governance guidance for Sure Start Children's Centres and extended schools. [Online] [Accessed 19 January 2010]. Available from https://www.education.gov.uk/

DEWALT, D. A. & HINK, A. 2009. Health literacy and child health outcomes: a systematic review of the literature. *Pediatrics,* 124 Suppl 3**,** S265-74.

DEX, S. & JOSHI, H. 2005. *Children of the 21st century: From birth to nine months.* Bristol: The Policy Press.

DIEGO, M. A., FIELD, T., HERNANDEZ-REIF, M., CULLEN, C., SCHANBERG, S. & KUHN, C. 2004. Prepartum, postpartum, and chronic depression effects on newborns. *Psychiatry-Interpersonal and Biological Processes,* 67**,** 63-80.

DIETZ, L. J., JENNINGS, K. D., KELLEY, S. A. & MARSHAL, M. 2009. Maternal Depression, Paternal Psychopathology, and Toddlers' Behavior Problems. *Journal of Clinical Child and Adolescent Psychology,* 38**,** 48-61.

DODGE, J. & SILVA, P. A. 1980. A Study of Mothers Health. *New Zealand Medical Journal,* 91**,** 352-354.

DOGRA, A., PARKIN, A., GALE, F. & FRAKE, C. 2002. *A multidisciplinary handbook of child and adolescent mental health for front-line professionals*. London: Jessica Kingsley Publishers

DOMINO, G., & DOMINO, M.L. 2006. Psychological Testing. New York: Cambridge University Press.

D'ONISE, K., LYNCH, J. W., SAWYER, M. G. & MCDERMOTT, R. A. 2010. Can preschool improve child health outcomes? A systematic review. *Soc Sci Med,* 70**,** 1423-40.

DOYLE, O., HARMON, C. P., HECKMAN, J. J. & TREMBLAY, R. E. 2009. Investing in early human development: timing and economic efficiency. *Econ Hum Biol,* 7**,** 1-6.

DUNCAN, G. J. & BROOKS-GUNN, J. 1997. *Consequences of growing up poor.,* New York, Russell Sage Foundation Press.

DUNCAN, G. J., DOWSETT, C. J., CLAESSENS, A., MAGNUSON, K., HUSTON, A. C., KLEBANOV, P., PAGANI, L. S., FEINSTEIN, L., ENGEL, M., BROOKS-GUNN, J., SEXTON, H., DUCKWORTH, K. & JAPEL, C. 2007. School readiness and later achievement. *Dev Psychol,* 43**,** 1428-46.

EAGLE, M. 1995. The developmental perspectives of attachment and psychoanalytic theory. In Goldberg, S. et al. (eds.). *Attachment Theory. Social, Developmental, and Clinical Perspectives.* London: The Analytic Press pp. 123-150

EATON, W. W., SHAO, H., NESTADT, G., LEE, H. B., BIENVENU, O. J. & ZANDI, P. 2008. Population-based study of first onset and chronicity in major depressive disorder. *Arch Gen Psychiatry,* 65**,** 513-20.

EBESUTANI, C., BERNSTEIN, A., NAKAMURA, B. J., CHORPITA, B. F., HIGA-MCMILLAN, C. K., WEISZ, J. R. & THE RESEARCH NETWORK ON YOUTH MENTAL, H. 2010. Concurrent Validity of the Child Behavior Checklist DSM-Oriented Scales: Correspondence with DSM Diagnoses and Comparison to Syndrome Scales. *J Psychopathol Behav Assess,* 32**,** 373-384.

EDGE, D., BAKER, D. & ROGERS, A. 2004. Perinatal depression among black Caribbean women. *Health Soc Care Community,* 12**,** 430-8.

EGELAND, B., KALKOSKE, M., GOTTESMAN, N. & ERICKSON, M. F. 1990. Preschool behavior problems: stability and factors accounting for change. *J Child Psychol Psychiatry,* 31**,** 891-909.

EISENBERG, N., CUMBERLAND, A., SPINRAD, T. L., FABES, R. A., SHEPARD, S. A., REISER, M., MURPHY, B. C., LOSOYA, S. H. & GUTHRIE, I. K. 2001. The relations of regulation and emotionality to children's externalizing and internalizing problem behavior. *Child Dev,* 72**,** 1112-34.

ELDER, G. H., JR. 1994. Time, Human Agency, and Social Change: Perspectives on the Life Course. *Social Psychology Quarterly,* 57**,** 4-15.

ELLIOT, C.D. 1983. *British Abilities Scale. Technical manual*. Windsor, Berkshire: NFER-Nelson.

ELLIOTT, C. D., MURRAY, D. & PEARSON, L. S. 1979. *British Ability Scales*. Windsor, England: National Foundation for Educational Research.

ESKENAZI, B. & BERGMANN, J. J. 1995. Passive and Active Maternal Smoking during Pregnancy, as Measured by Serum Cotinine, and Postnatal Smoke Exposure .1.

Effects on Physical Growth at Age 5 Years. *American Journal of Epidemiology,* 142**,** S10-S18.

ESSEX, M. J., KLEIN, M. H., CHO, E. & KALIN, N. H. 2002. Maternal stress beginning in infancy may sensitize children to later stress exposure: Effects on cortisol and behavior. *Biological Psychiatry,* 52**,** 776-784.

FAIRCLOUGH, D.L. 2002. *Design and Analysis of Quality of Life Studies in Clinical Trials: Interdisciplinary Statistics.* New York: Chapman & Hall/CRC

FAVA, G. A., PARK, S. K. & SONINO, N. 2006. Treatment of recurrent depression. *Expert Rev Neurother,* 6**,** 1735-40.

FAWCETT, S. L., WANG, Y. Z. & BIRCH, E. E. 2005. The critical period for susceptibility of human stereopsis. *Invest Ophthalmol Vis Sci,* 46**,** 521-5.

FEINSTEIN, L. 2003. Inequality in the early cognitive development of British children in the 1970 Cohort. *Economica,* 70**,** 73-97.

FEINSTEIN, L. & BYNNER, J. 2004. The importance of cognitive development in middle childhood for adulthood socioeconomic status, mental health, and problem behavior. *Child Dev,* 75**,** 1329-39.

FELDMAN, R. 2007. Parent-infant synchrony and the construction of shared timing; physiological precursors, developmental outcomes, and risk conditions. *J Child Psychol Psychiatry,* 48**,** 329-54.

FERGUSSON, D. M., JOHN HORWOOD, L. & RIDDER, E. M. 2005. Show me the child at seven: the consequences of conduct problems in childhood for psychosocial functioning in adulthood. *Journal of Child Psychology and Psychiatry,* 46**,** 837-849.

FIELD, A. P. (2009) *Discovering statistics using SPSS (3rd Edition).* London: Sage pp783

FIELD, T. 2010. Postpartum depression effects on early interactions, parenting, and safety practices: A review. *Infant Behavior & Development,* 33**,** 1-6.

FIHRER, I., MCMAHON, C. A. & TAYLOR, A. J. 2009. The impact of postnatal and concurrent maternal depression on child behaviour during the early school years. *Journal of Affective Disorders,* 119**,** 116-123.

FISCHER, M., ROLF, J. E., HASAZI, J. E. & CUMMINGS, L. 1984. Follow-up of a preschool epidemiological sample: cross-age continuities and predictions of later adjustment with internalizing and externalizing dimensions of behavior. *Child Dev,* 55**,** 137-50.

FRAZIER, T. W., YOUNGSTROM, E. A., GLUTTING, J. J. & WATKINS, M. W. 2007. ADHD and achievement: meta-analysis of the child, adolescent, and adult literatures and a concomitant study with college students. *J Learn Disabil,* 40**,** 49-65.

GALOBARDES, B., LYNCH, J. W. & DAVEY SMITH, G. 2004. Childhood socioeconomic circumstances and cause-specific mortality in adulthood: systematic review and interpretation. *Epidemiol Rev,* 26**,** 7-21.

GALOBARDES, B., LYNCH, J. W. & DAVEY SMITH, G. 2007. Is the association between childhood socioeconomic circumstances and cause-specific mortality established? Update of a systematic review. . *Journal of Epidemiology and Community Health,* 62**,** 387-390.

GARCES, E., DUNCAN, T. & CURRIE, J. 2002. Longer-Term Effects of Head Start. *American Economic Review, 92,* 999-1012.

GASMAN, I., PURPER-OUAKIL, D., MICHEL, G., MOUREN-SIMEONI, M. C., BOUVARD, M., PEREZ-DIAZ, F. & JOUVENT, R. 2002. Cross-cultural assessment of childhood temperament - A confirmatory factor analysis of the French Emotionality Activity and Sociability (EAS) Questionnaire. *European Child & Adolescent Psychiatry,* 11**,** 101-107.

GATHERCOLE, S. E., WILLIS, C. S., EMSLIE, H. & BADDELEY, A. D. 1992. Phonological Memory and Vocabulary Development during the Early School Years - a Longitudinal-Study. *Developmental Psychology,* 28**,** 887-898.

GEDDES, J. R., CARNEY, S. M., DAVIES, C., FURUKAWA, T. A., KUPFER, D. J., FRANK, E. & GOODWIN, G. M. 2003. Relapse prevention with antidepressant drug treatment in depressive disorders: a systematic review. *Lancet,* 361**,** 653-61.

GEDDES, R., HAW, S. & FRANK, J. 2010. *Interventions for Promoting Early Child Development for Health.* [Online]. [Accessed 24th June 2011].

GELFAND, D. M. & TETI, D. M. 1990. The Effects of Maternal Depression on Children. *Clinical Psychology Review,* 10**,** 329-353.

GENNETIAN, L., & MILLER, C. 2000. *Reforming welfare and rewarding work: Final report of the Minnesota Family Investment Program: Vol. 2. Effects on children*. New York: Manpower Demonstration Research Corporation.

GERHARDS, S. A., DE GRAAF, L. E., JACOBS, L. E., SEVERENS, J. L., HUIBERS, M. J., ARNTZ, A., RIPER, H., WIDDERSHOVEN, G., METSEMAKERS, J. F. & EVERS, S. M. 2010. Economic evaluation of online computerised cognitive-behavioural therapy without support for depression in primary care: randomised trial. *Br J Psychiatry,* 196**,** 310-8.

GINSBURG, K. R., AMERICAN ACADEMY OF PEDIATRICS COMMITTEE ON, C., AMERICAN ACADEMY OF PEDIATRICS COMMITTEE ON PSYCHOSOCIAL ASPECTS OF, C. & FAMILY, H. 2007. The importance of play in promoting healthy child development and maintaining strong parent-child bonds. *Pediatrics,* 119**,** 182-91.

GJERDINGEN, D. 2003. The effectiveness of various postpartum depression treatments and the impact of antidepressant drugs on nursing infants. *J Am Board Fam Pract,* 16**,** 372-82.

GLASCOE, F. P. 2005. Screening for developmental and behavioral problems. *Ment Retard Dev Disabil Res Rev,* 11**,** 173-9.

GLASS, N. 1999. Sure Start: The Development of an Early Intervention Programme for Young Children in the United Kingdom. *Children and Society,* 13**,** 257-264.

GOLDMAN, J., YU, M., SALGANIK, D. T., SCOTT, F. C., BENNETT, L. L. & KAY, J. C. 2006. Early Intervention in Low Birth Weight Premature Infants: Results at 18 Years of Age for the Infant Health and Development Program. *Pediatrics,* 117**,** 771-780.

GOLDWYN, R., STANLEY, C., SMITH, V. & GREEN, J. 2000. The Manchester Child Attachment Story Task: relationship with parental AAI, SAT and child behaviour. *Attach Hum Dev,* 2**,** 71-84.

GOODMAN, A. & GOODMAN, R. 2009. Strengths and difficulties questionnaire as a dimensional measure of child mental health. *J Am Acad Child Adolesc Psychiatry,* 48**,** 400-3.

GOODMAN, R. 1997. The Strengths and Difficulties Questionnaire: a research note. *J Child Psychol Psychiatry,* 38**,** 581-6.

GOODMAN, R. 2001. Psychometric properties of the strengths and difficulties questionnaire. *Journal of the American Academy of Child and Adolescent Psychiatry,* 40**,** 1337-1345.

GOODMAN, S. H. & GOTLIB, I. H. 1999. Risk for psychopathology in the children of depressed mothers: a developmental model for understanding mechanisms of transmission. *Psychol Rev,* 106**,** 458-90.

GOODMAN, S. H., ROUSE, M. H., CONNELL, A. M., BROTH, M. R., HALL, C. M. & HEYWARD, D. 2011. Maternal Depression and Child Psychopathology: A Meta-Analytic Review. *Clinical Child and Family Psychology Review,* 14**,** 1-27.

GORELICK, M. H. 2006. Bias arising from missing data in predictive models. *Journal of Clinical Epidemiology,* 59**,** 1115-1123.

GRACE, S. L., EVINDAR, A. & STEWART, D. E. 2003. The effect of postpartum depression on child cognitive development and behavior: a review and critical analysis of the literature. *Arch Womens Ment Health,* 6**,** 263-74.

GRAN, B. 1995. Major differences in cardiovascular risk indicators by educational status. Results from a population based screening program. *Scand J Soc Med,* 23**,** 9-16.

GRANOT, D. & MAYSELESS, O. 2001. Attachment security and adjustment to school in middle childhood. *International Journal of Behavioral Development,* 25**,** 530-541.

GRANT, G., NOLAN, M. & ELLIS, N. 1990. A reappraisal of the Malaise Inventory. *Soc Psychiatry Psychiatr Epidemiol,* 25**,** 170-8.

GREEN, J., STANLEY, C., SMITH, V. & GOLDWYN, R. 2000. A new method of evaluating attachment representations in young school-age children: the Manchester Child Attachment Story Task. *Attach Hum Dev,* 2**,** 48-70.

GREGOIRE, A. J., KUMAR, R., EVERITT, B., HENDERSON, A. F. & STUDD, J. W. 1996. Transdermal oestrogen for treatment of severe postnatal depression. *Lancet,* 347**,** 930-3.

GRIMES, D. A. & SCHULZ, K. F. 2002. Uses and abuses of screening tests. *Lancet,* 359**,** 881-4.

GRYCH, J. H. & FINCHAM, F. D. (Eds) 2001. *Interparental conflict and child development: theory, research, and applications*. Cambridge, MA: Cambridge University Press.

GUTTMANNOVA, K., SZANYI, J. M. & CALI, P. W. 2008. Internalizing and externalizing behavior problem scores - Cross-ethnic and longitudinal measurement

invariance of the Behavior Problem Index. *Educational and Psychological Measurement,* 68**,** 676-694.

HALBREICH, U. & KARKUN, S. 2006. Cross-cultural and social diversity of prevalence of postpartum depression and depressive symptoms. *Journal of Affective Disorders,* 91**,** 97-111.

HALLAM, S. 2012. Streaming and Setting in UK Primary Schools: evidence from the Millennium Cohort Study. *Forum,* 54**,** 57-64.

HALONEN, A., AUNOLA, K., AHONEN, T. & NURMI, J. E. 2006. The role of learning to read in the development of problem behaviour: a cross-lagged longitudinal study. *Br J Educ Psychol,* 76**,** 517-34.

HAMILTON, H. A. & HANGOW, D. W. 2008. Symptoms of Psychological Distress Among Mothers and their Adult Offspring: A Longitudinal Examination. *Journal of Child and Family Studies,* 17.

HAMMEN, C. & BRENNAN, P. A. 2003. Severity, chronicity, and timing of maternal depression and risk for adolescent offspring diagnoses in a community sample. *Arch Gen Psychiatry,* 60**,** 253-8.

HAN, W. J., WALDFOGEL, J. & BROOKS-GUNN, J. 2001. The effects of early maternal employment on later cognitive and behavioral outcomes. *Journal of Marriage and Family,* 63**,** 336-354.

HANSEN, K., JOSHI, H. & DEX, S. 2010. *Children of the 21st century: the first five years. ,* Bristol, Policy Press.

HANSEN, K., JONES, E., JOSHI, H., & BUDGE, D. 2010. *Millennium Cohort Study Fourth Survey: A User's Guide to Initial Findings. 2nd edition.* London: Centre for Longitudinal Studies

HARBERGER, P. N., BERCHOLD, N. G. & HONIKMAN, J. I. 1992. Cries for help. *In:* HAMILTON, J. A. & HARBINGER, P. N. (eds.) *Postpartum Depression: A Picture Puzzle.* . Philadelphia: University of Pennsylvania Press.

HARRIS, B., LOVETT, L., SMITH, J., READ, G., WALKER, R. & NEWCOMBE, R. 1996. Cardiff puerperal mood and hormone study. III. Postnatal depression at 5 to 6 weeks postpartum, and its hormonal correlates across the peripartum period. *Br J Psychiatry,* 168**,** 739-44.

HAY, D. F. & KUMAR, R. 1995. Interpreting the effects of mothers' postnatal depression on children's intelligence: a critique and re-analysis. *Child Psychiatry Hum Dev,* 25**,** 165-81.

HAY, D. F., PAWLBY, S., SHARP, D., ASTEN, P., MILLS, A. & KUMAR, R. 2001. Intellectual problems shown by 11-year-old children whose mothers had postnatal depression. *J Child Psychol Psychiatry,* 42**,** 871-89.

HECKMAN, J. J. 2006. Skill formation and the economics of investing in disadvantaged children. *Science,* 312**,** 1900-2.

HECKMAN, J. J. 2007. The economics, technology, and neuroscience of human capability formation. *Proc Natl Acad Sci U S A,* 104**,** 13250-5.

HECKMAN, J. J. 2008. Schools, Skills, and Synapses. *Econ Inq,* 46**,** 289.

HECKMAN, J.J. Invest in the very young. 2nd edition. In TREMBLAY R.E., BOIVIN, M., PETERS, R.D., eds. *Encyclopedia on Early Childhood Development [online]*. Montreal, Quebec: Centre of Excellence for Early Childhood Development and Strategic Knowledge Cluster on Early Child Development; 2007:1-2. Available at: http://www.child-encyclopedia.com/documents/HeckmanANGxp.pdf. Accessed [10th January 2011].

HECKMAN, J.J., 2008. The case for investing in disadvantaged children. In: First Focus (Ed.), *Big Ideas for Children: Investing in Our Nation's Future*. Washington, DC: First Focus, pp. 49–58.

HEALEY, A., KNAPP, M. & FARRINGTON, D. P. 2004. Adult labour market implications of antisocial behaviour in childhood and adolescence: findings from a UK longitudinal study. *Applied Economics,* 36**,** 93-105.

HEARN, G., ILIFF, A., JONES, I., KIRBY, A., ORMISTON, P., PARR, P., ROUT, J. & WARDMAN, L. 1998. Postnatal depression in the community. *Br J Gen Pract,* 48**,** 1064-6.

HECKMAN, J. J. 2006. Skill formation and the economics of investing in disadvantaged children. *Science,* 312**,** 1900-2.

HENDRICK, V. 2003. Treatment of postnatal depression. *BMJ,* 327**,** 1003-4.

HERTZMAN, C., POWER, C., MATTHEWS, S. & MANOR, O. 2001. Using an interactive framework of society and lifecourse to explain self-rated health in early adulthood. *Social Science & Medicine,* 53**,** 1575-1585.

HETHERINGTON, E. M. & JODL, K. M. 1994. Stepfamilies as Settings for Child Development. *Stepfamilies***,** 55-79.

HETHERINGTON, E. M., PARKE, R.D., GAUVAIN, M. & LOCKE, V.O. 2006. Child Psychology: A Contemporary Viewpoint. London: McGraw-Hill pp. 238-240.

HEWITT, C., GILBODY, S., BREALEY, S., PAULDEN, M., PALMER, S., MANN, R., GREEN, J., MORRELL, J., BARKHAM, M., LIGHT, K. & RICHARDS, D. 2009. Methods to identify postnatal depression in primary care: an integrated evidence synthesis and value of information analysis. *Health Technol Assess,* 13**,** 1-145, 147-230.

HEWITT, C. E. & GILBODY, S. M. 2009. Is it clinically and cost effective to screen for postnatal depression: a systematic review of controlled clinical trials and economic evidence. *BJOG,* 116**,** 1019-27.

HILL, V. 2005. Through the past darkly: A review of the British ability scales second edition. *Child and Adolescent Mental Health, 10*, 87–98.

HINSHAW, S. P. 1987. On the distinction between attentional deficits/hyperactivity and conduct problems/aggression in child psychopathology. *Psychol Bull,* 101**,** 443-63.

HIRST, M. A. & BRADSHAW, J. R. 1983. Evaluating the Malaise Inventory - a Comparison of Measures of Stress. *Journal of Psychosomatic Research,* 27**,** 193-199.

HJELMGREN, J., BERGGREN, F. & ANDERSSON, F. 2001. Health economic guidelines--similarities, differences and some implications. *Value Health,* 4**,** 225-50.

HO, C., BLUESTEIN, D. N. & JENKINS, J. M. 2008. Cultural differences in the relationship between parenting and children's behavior. *Developmental Psychology,* 44**,** 507-522.

HOBCRAFT, J. & KIERNAN, K. (2010) *Predictive factors from age 3 and infancy for poor child outcomes at age 5 relating to children's development, behaviour and health: evidence from the Millennium Cohort Study: Report prepared for The National child and Maternal Health Observatory (ChiMat) on behalf of the Department of Health.* University of York.

HODDINOTT, J., BEHRMAN, J. R., MALUCCIO, J. A., MELGAR, P., QUISUMBING, A. R., RAMIREZ-ZEA, M., STEIN, A. D., YOUNT, K. M. & MARTORELL, R. 2013. Adult consequences of growth failure in early childhood. *Am J Clin Nutr,* 98**,** 1170-8.

HOEYMANS, N., SMIT, H. A., VERKLEIJ, H. & KROMHOUT, D. 1996. Cardiovascular risk factors in relation to educational level in 36 000 men and women in The Netherlands. *Eur Heart J,* 17**,** 518-25.

HOFFMAN, L., & YOUNGBLADE, L. 1999. *Mothers at work: Effects on children's wellbeing*. New York: Cambridge University Press.

HOLMSHAW, J. & SIMONOFF, E. 1996. Retrospective recall of childhood psychopathology. *International Journal of Methods in Psychiatric Research,* 6**,** 79-88.

HOLOPAINEN, D. 2002. The experience of seeking help for postnatal depression. *Aust J Adv Nurs,* 19**,** 39-44.

HONEY, K. L., BENNETT, P. & MORGAN, M. 2003. Predicting postnatal depression. *J Affect Disord,* 76**,** 201-10.

HORWITZ, S. M., BRIGGS-GOWAN, M. J., STORFER-ISSER, A. & CARTER, A. S. 2007. Prevalence, correlates, and persistence of maternal depression. *J Womens Health (Larchmt),* 16**,** 678-91.

HORWITZ, S. M., IRWIN, J. R., BRIGGS-GOWAN, M. J., HEENAN, J. M. B., MENDOZA, J. & CARTER, A. S. 2003. Language delay in a community cohort of young children. *Journal of the American Academy of Child and Adolescent Psychiatry,* 42**,** 932-940.

HORWOOD, L. J., MOGRIDGE, N. & DARLOW, B. A. 1998. Cognitive, educational, and behavioural outcomes at 7 to 8 years in a national very low birthweight cohort. *Arch Dis Child Fetal Neonatal Ed,* 79**,** F12-20.

HOULE, T. T. & STUMP, D. A. 2008. Statistical significance versus clinical significance. *Semin Cardiothorac Vasc Anesth,* 12**,** 5-6.

HOWARD, K., MARTIN, A., BERLIN, L. J. & BROOKS-GUNN, J. 2011. Early mother-child separation, parenting, and child well-being in Early Head Start families. *Attach Hum Dev,* 13**,** 5-26.

HUNINK, M., GLAZIOU, P., SIEGEL, J., WEEKS, J., PLISKIN, J., ELSTEIN, A. & WEINSTEIN, M. 2001. *Decision making in health and medicine. Integrating evidence and values.,* Cambridge, Cambridge University Press.

INGRAM, R. E. & RITTER, J. 2000. Vulnerability to depression: cognitive reactivity and parental bonding in high-risk individuals. *J Abnorm Psychol,* 109**,** 588-96.

IRIBARREN, C., LUEPKER, R. V., MCGOVERN, P. G., ARNETT, D. K. & BLACKBURN, H. 1997. Twelve-year trends in cardiovascular disease risk factors in the Minnesota Heart Survey. Are socioeconomic differences widening? *Arch Intern Med,* 157**,** 873-81.

JACKSON, C., TROWER, P., REID, I., SMITH, J., HALL, M., TOWNEND, M., BARTON, K., JONES, J., ROSS, K., RUSSELL, R., NEWTON, E., DUNN, G. & BIRCHWOOD, M. 2009. Improving psychological adjustment following a first episode of psychosis: a randomised controlled trial of cognitive therapy to reduce post psychotic trauma symptoms. *Behav Res Ther,* 47**,** 454-62.

JELLEYMAN, T. & SPENCER, N. 2008. Residential mobility in childhood and health outcomes: a systematic review. *J Epidemiol Community Health,* 62**,** 584-92.

JOSEFSSON, A. & SYDSJO, G. 2007. A follow-up study of postpartum depressed women: recurrent maternal depressive symptoms and child behavior after four years. *Archives of Womens Mental Health,* 10**,** 141-145.

JOSHI, R.M. 2005. Vocabulary: A critical component of comprehension. *Reading & Writing Quarterly,* 21, 209-219.

KAESTNER, R. & CORMAN, H. 199). *The Impact of Child Health and Fmaily Inputs on Child Cognitive Development*. NBER Working Paper #5257, Cambridge, MA: NBER

KALTENTHALER, E., SHACKLEY, P., STEVENS, K., BEVERLEY, C., PARRY, G. & CHILCOTT, J. 2002. A systematic review and economic evaluation of computerised cognitive behaviour therapy for depression and anxiety. *Health Technol Assess,* 6**,** 1-89.

KALTENTHALER, E., PARRY, G., BEVERLEY, C. & FERRITER, M. 2008. Computerised cognitive-behavioural therapy for depression: systematic review. *Br J Psychiatry,* 193**,** 181-4.

KAROLY, L.A., KILBURN, M.R. & CANNON, J.S. 2005. *Early childhood interventions: Proven results, future promise*. Santa Monica, CA: RAND Corporation

KATTAN, M.W. & COWEN, M.E. 2009. *Encyclopaedia of Medical Decision Making*. London: Sage pp 499

KAUFMAN, A. S., & LICHTENBERGER, E. O. (2000). *Essentials of WISC-III and WPPSI-R assessment*. John Wiley & Sons Inc.

KENWARD, M. G. & CARPENTER, J. 2007. Multiple imputation: current perspectives. *Statistical Methods in Medical Research,* 16**,** 199-218.

KENDALL-TACKETT, K. A. 2009. *Depression in new mothers: Causes, consequences, and treatment alternatives*. Routledge.

KERSTEN-ALVAREZ, L. E., HOSMAN, C. M., RIKSEN-WALRAVEN, J. M., VAN DOESUM, K. T., SMEEKENS, S. & HOEFNAGELS, C. 2012. Early school outcomes for children of postpartum depressed mothers: comparison with a community sample. *Child Psychiatry Hum Dev,* 43**,** 201-18.

KESSLER, R. C., ANDREWS, G., COLPE, L. J., HIRIPI, E., MROCZEK, D. K.,
NORMAND, S. L., WALTERS, E. E. & ZASLAVSKY, A. M. 2002. Short
screening scales to monitor population prevalences and trends in non-specific
psychological distress. *Psychol Med,* 32**,** 959-76.

KESSLER, R. C., BARKER, P. R., COLPE, L. J., EPSTEIN, J. F., GFROERER, J. C.,
HIRIPI, E., HOWES, M. J., NORMAND, S. L., MANDERSCHEID, R. W.,
WALTERS, E. E. & ZASLAVSKY, A. M. 2003. Screening for serious mental
illness in the general population. *Arch Gen Psychiatry,* 60**,** 184-9.

KESSLER, R. C., GREEN, J. G., GRUBER, M. J., SAMPSON, N. A., BROMET, E.,
CUITAN, M., FURUKAWA, T. A., GUREJE, O., HINKOV, H., HU, C. Y.,
LARA, C., LEE, S., MNEIMNEH, Z., MYER, L., OAKLEY-BROWNE, M.,
POSADA-VILLA, J., SAGAR, R., VIANA, M. C. & ZASLAVSKY, A. M. 2010.
Screening for serious mental illness in the general population with the K6 screening
scale: results from the WHO World Mental Health (WMH) survey initiative. *Int J
Methods Psychiatr Res,* 19 Suppl 1**,** 4-22.

KIERNAN, K. E. & HUERTA, M. C. 2008. Economic deprivation, maternal depression,
parenting and children's cognitive and emotional development in early childhood.
British Journal of Sociology, 59**,** 783-806.

KIERNAN, K. E. & MENSAH, F. K. 2009. Poverty, Maternal Depression, Family Status
and Children's Cognitive and Behavioural Development in Early Childhood: A
Longitudinal Study. *Journal of Social Policy,* 38**,** 569-588.

KIERNAN, K.E. AND MENSAH, F.K. 2009. *Maternal indicators in pregnancy and
children's infancy that signal future outcomes for children's development,
behaviour and health: evidence from the Millennium Cohort Study*. University of
York

KINCAID, C. 2005. Guidelines for selecting the covariance structure in mixed model
analysis. *Statistics and Data Analysis, 30,* 1-8.

KITZMAN, H., OLDS, D. L., HENDERSON, C. R., JR., HANKS, C., COLE, R., TATELBAUM, R., MCCONNOCHIE, K. M., SIDORA, K., LUCKEY, D. W., SHAVER, D., ENGELHARDT, K., JAMES, D. & BARNARD, K. 1997. Effect of prenatal and infancy home visitation by nurses on pregnancy outcomes, childhood injuries, and repeated childbearing. A randomized controlled trial. *JAMA,* 278**,** 644-52.

KLEIN, R. G., MANNUZZA, S., OLAZAGASTI, M. A., ROIZEN, E., HUTCHISON, J. A., LASHUA, E. C. & CASTELLANOS, F. X. 2012. Clinical and functional outcome of childhood attention-deficit/hyperactivity disorder 33 years later. *Arch Gen Psychiatry,* 69**,** 1295-303.

KNAPP, M., KING, D., HEALEY, A. & THOMAS, C. 2011. Economic outcomes in adulthood and their associations with antisocial conduct, attention deficit and anxiety problems in childhood. *J Ment Health Policy Econ,* 14**,** 137-47.

KNUDSEN, E. I., HECKMAN, J. J., CAMERON, J. L. & SHONKOFF, J. P. 2006. Economic, neurobiological, and behavioral perspectives on building America's future workforce. *Proc Natl Acad Sci U S A,* 103**,** 10155-62.

KORHONEN, M., LUOMA, I., SALMELIN, R. & TAMMINEN, T. 2012. A longitudinal study of maternal prenatal, postnatal and concurrent depressive symptoms and adolescent well-being. *Journal of Affective Disorders,* 136**,** 680-692.

KOTULAK, R. 1998. Inside the brain: revolutionary discoveries of how the mind works. *Prev Med,* 27**,** 246-7.

KROENKE, K., SPITZER, R. L. & WILLIAMS, J. B. 2001. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med,* 16**,** 606-13.

KUH, D. & BEN-SHLOMO, Y. 1997. *A life course approach to chronic disease epidemiology. ,* Oxford, Oxford University Press.

KUH, D., BEN-SHLOMO, Y., LYNCH, J., HALLQVIST, J. & POWER, C. 2003. Life course epidemiology. *Journal of Epidemiology and Community Health,* 57**,** 778-783.

LAMBERT, N. M. 1988. Adolescent outcomes for hyperactive children. Perspectives on general and specific patterns of childhood risk for adolescent educational, social, and mental health problems. *Am Psychol,* 43**,** 786-99.

LAMBERT, N. M., HARTSOUGH, C. S. & SANDOVAL, J. 1990. *Manual for the Children's Attention and Adjustment Survey.,* Palo Alto, CA, Consulting Psychologists Press.

LEE S, AOS S, DRAKE E, PENNUCCI A, MILLER M, ANDERSON L. 2012 Return on investment: Evidence-based options to improve state-wide outcomes, April 2012 (Document No. 12-04-1201) Olympia: Washington State Institute for Public Policy.

LEMCHE, E., KREPPNER, J. M., JORASCHKY, P. & KLANN-DELIUS, G. 2007. Attachment organization and the early development of internal state language: A longitudinal perspective. *International Journal of Behavioral Development,* 31**,** 252-262.

LIAW, F. & BROOKS-GUNN, J. 199). Cumulative familiar risks and low-birthweight children's cognitive and behavioral development. *Journal of Clinical Child Psychology, 23,* 360-372.

LITTLE, M., BJORNSTAD, G., & EDOVALD, T. 2013. Investing in Children: An Overview. Social Research Unit: Dartington (UK).

LOVEJOY, M. C., GRACZYK, P. A., O'HARE, E. & NEUMAN, G. 2000. Maternal depression and parenting behavior: a meta-analytic review. *Clin Psychol Rev,* 20**,** 561-92.

LUOMA, I., TAMMINEN, T., KAUKONEN, P., LAIPPALA, P., PUURA, K., SALMELIN, R. & ALMQVIST, F. 2001. Longitudinal study of maternal depressive symptoms and child well-being. *J Am Acad Child Adolesc Psychiatry,* 40**,** 1367-74.

LYONS-RUTH, K. 1996. Attachment relationships among children with aggressive behavior problems: the role of disorganized early attachment patterns. *J Consult Clin Psychol,* 64**,** 64-73.

LYONS-RUTH, K., BRONFMAN, E. & PARSONS, E. 1999. Atypical attachment in infancy and early childhood among children at developmental risk. IV. Maternal frightened, frightening, or atypical behavior and disorganized infant attachment patterns. *Monogr Soc Res Child Dev,* 64**,** 67-96; discussion 213-20.

MACRAN, S., CLARKE, L. & JOSHI, H. 1996. Women's health: dimensions and differentials. *Soc Sci Med,* 42**,** 1203-16.

MADIGAN, S., ATKINSON, L., LAURIN, K. & BENOIT, D. 2013. Attachment and Internalizing Behavior in Early Childhood: A Meta-Analysis. *Developmental Psychology,* 49**,** 672-689.

MADIGAN, S., MORAN, G., SCHUENGEL, C., PEDERSON, D. R. & OTTEN, R. 2007. Unresolved maternal attachment representations, disrupted maternal behavior and disorganized attachment in infancy: links to toddler behavior problems. *Journal of Child Psychology and Psychiatry,* 48**,** 1042-1050.

MANN, R., ADAMSON, J. & GILBODY, S. M. 2012. Diagnostic accuracy of case-finding questions to identify perinatal depression. *Canadian Medical Association Journal,* 184**,** E424-E430.

MANN, S. L., WADSWORTH, M. E. & COLLEY, J. R. 1992. Accumulation of factors influencing respiratory illness in members of a national birth cohort and their offspring. *J Epidemiol Community Health,* 46**,** 286-92.

MANNING, W. D. & LAMB, K. A. 2003. Adolescent well-being in cohabiting, married, and single-parent families. *Journal of Marriage and Family,* 65**,** 876-893.

MARCH, J., SILVA, S., VITIELLO, B. & TEAM, T. 2006. The Treatment for Adolescents with Depression Study (TADS): methods and message at 12 weeks. *J Am Acad Child Adolesc Psychiatry,* 45**,** 1393-403.

MARTINS, C. & GAFFAN, E. A. 2000. Effects of early maternal depression on patterns of infant-mother attachment: A meta-analytic investigation. *Journal of Child Psychology and Psychiatry and Allied Disciplines,* 41**,** 737-746.

MASSETTI, G. M., LAHEY, B. B., PELHAM, W. E., LONEY, J., EHRHARDT, A., LEE, S. S. & KIPP, H. 2008. Academic achievement over 8 years among children who met modified criteria for attention-deficit/hyperactivity disorder at 4-6 years of age. *J Abnorm Child Psychol,* 36**,** 399-410.

MATTHEY, S., WHITE, T., PHILLIPS, J., TAOUK, R., CHEE, T. T. & BARNETT, B. 2005. Acceptability of routine antenatal psychosocial assessments to women from English and non-English speaking backgrounds. *Arch Womens Ment Health,* 8**,** 171-80.

MAYS, N., ROBERTS, E. & POPAY, J. 2001. Synthesising research evidence. . *In:* FULOP, N., ALLEN, P., CLARKE, A. & BLACK, N. (eds.) *Studying the Organisation and Delivery of Health Services: Research Methods.* London: Routledge.

MCCRONE, P., KNAPP, M., PROUDFOOT, J., RYDEN, C., CAVANAGH, K., SHAPIRO, D. A., ILSON, S., GRAY, J. A., GOLDBERG, D., MANN, A., MARKS, I., EVERITT, B. & TYLEE, A. 2004. Cost-effectiveness of computerised cognitive-behavioural therapy for anxiety and depression in primary care: randomised controlled trial. *Br J Psychiatry,* 185**,** 55-62.

MCGEE, R., PARTRIDGE, F., WILLIAMS, S. & SILVA, P. A. 1991. A twelve-year follow-up of preschool hyperactive children. *J Am Acad Child Adolesc Psychiatry,* 30**,** 224-32.

MCGEE, R., WILLIAMS, S. & SILVA, P. A. 1986. An evaluation of the Malaise Inventory. *J Psychosom Res,* 30**,** 147-52.

MCINTOSH, J. 1993. Postpartum depression: women's help-seeking behaviour and perceptions of cause. *J Adv Nurs,* 18**,** 178-84.

MCKAY, S. 2003. The Dynamics of Lone Parents, Employment and Poverty in Great Britain. *Sociology and Social Politics, 6,* 101-124

MCLENNAN, J. D., KOTELCHUCK, M. & CHO, H. 2001. Prevalence, persistence, and correlates of depressive symptoms in a national sample of mothers of toddlers. *J Am Acad Child Adolesc Psychiatry,* 40**,** 1316-23.

MCLEOD, J. D. & KAISER, K. 2004. Childhood emotional and behavioural problems and educational attainment. *American Sociological Review***,** 636-658.

MCLOYD, V. C. 1998. Socioeconomic disadvantage and child development. *Am Psychol,* 53**,** 185-204.

MCMAHON, C., BARNETT, B., KOWALENKO, N. & TENNANT, C. 2005. Psychological factors associated with persistent postnatal depression: past and current relationships, defence styles and the mediating role of insecure attachment style. *J Affect Disord,* 84**,** 15-24.

MELCHIOR, M., CHASTANG, J. F., DE LAUZON, B., GALERA, C., SAUREL-CUBIZOLLES, M. J., LARROQUE, B. & GRP, E. M.-C. C. S. 2012. Maternal depression, socioeconomic position, and temperament in early childhood: The EDEN mother-child cohort. *Journal of Affective Disorders,* 137**,** 165-169.

MENSAH, F. K. & KIERNAN, K. E. 2010. Parents' mental health and children's cognitive and social development. *Social Psychiatry and Psychiatric Epidemiology,* 45**,** 1023-1035.

MILGROM, J., WESTLEY, D. T. & GERNMILL, A. W. 2004. The mediating role of maternal responsiveness in some longer term effects of postnatal depression on infant development. *Infant Behavior & Development,* 27**,** 443-454.

MILGROM, J., GEMMILL, A. W., BILSZTA, J. L., HAYES, B., BARNETT, B., BROOKS, J., ERICKSEN, J., ELLWOOD, D. & BUIST, A. 2008. Antenatal risk factors for postnatal depression: a large prospective study. *J Affect Disord,* 108**,** 147-57.

MILGROM, J., NEGRI, L. M., GEMMILL, A. W., MCNEIL, M. & MARTIN, P. R. 2005. A randomized controlled trial of psychological interventions for postnatal depression. *Br J Clin Psychol,* 44**,** 529-42.

MISRI, S. & KENDRICK, K. 2008. Perinatal depression, fetal bonding, and mother-child attachment: a review of the literature. *Current Pediatric Review, 4,* 66-70.

MONTGOMERY, S. M., BARTLEY, M. J., COOK, D. G. & WADSWORTH, M. E. J. 1996. Health and social precursors of unemployment in young men in Great Britain. *Journal of Epidemiology and Community Health,* 50**,** 415-422.

MORRELL, J. & MURRAY, L. 2003. Parenting and the development of conduct disorder and hyperactive symptoms in childhood: a prospective longitudinal study from 2 months to 8 years. *Journal of Child Psychology and Psychiatry,* 44**,** 489-508.

MORRIS, P., & MICHALOPOULOS, C. 2000. *The Self Sufficiency Project at 36 Months: Effects on Children of a Program that Increased Employment and Income.* Ottawa: SRDC.

MURRAY, L. 1992. The impact of postnatal depression on infant development. *J Child Psychol Psychiatry,* 33**,** 543-61.

MURRAY, L., ARTECHE, A., FEARON, P., HALLIGAN, S., CROUDACE, T. & COOPER, P. 2010. The effects of maternal postnatal depression and child sex on academic performance at age 16 years: a developmental approach. *J Child Psychol Psychiatry,* 51**,** 1150-9.

MURRAY, L., ARTECHE, A., FEARON, P., HALLIGAN, S., GOODYER, I. & COOPER, P. 2011. Maternal postnatal depression and the development of depression in offspring up to 16 years of age. *J Am Acad Child Adolesc Psychiatry,* 50**,** 460-70.

MURRAY, L. & COOPER, P. 1997. Effects of postnatal depression on infant development. *Arch Dis Child,* 77**,** 99-101.

MURRAY, L. & COOPER, P. J. 1997. Postpartum depression and child development. *Psychol Med,* 27**,** 253-60.

MURRAY, L., COOPER, P. J., WILSON, A. & ROMANIUK, H. 2003. Controlled trial of the short- and long-term effect of psychological treatment of post-partum depression 2. Impact on the mother-child relationship and child outcome. *British Journal of Psychiatry,* 182**,** 420-427.

MURRAY, L., SINCLAIR, D., COOPER, P., DUCOURNAU, P. & TURNER, P. 1999. The socioemotional development of 5-year-old children of postnatally depressed mothers. *Journal of Child Psychology and Psychiatry,* 40**,** 1259-1271.

MYERS, W. R. 2000. Handling missing data in clinical trials: An overview. *Drug Information Journal,* 34**,** 525-533.

NAGATA, M., NAGAI, Y., SOBAJIMA, H., ANDO, T. & HONJO, S. 2003. Depression in the mother and maternal attachment - Results from a follow-up study at 1 year postpartum. *Psychopathology,* 36**,** 142-151.

NAHAS, V. L., HILLEGE, S. & AMASHEH, N. 1999. Postpartum depression. The lived experiences of Middle Eastern migrant women in Australia. *J Nurse Midwifery,* 44**,** 65-74.

NAJMAN, J. M., AIRD, R., BOR, W., O'CALLAGHAN, M., WILLIAMS, G. M. & SHUTTLEWOOD, G. J. 2004. The generational transmission of socioeconomic inequalities in child cognitive development and emotional health. *Soc Sci Med,* 58**,** 1147-58.

NATIONAL INSTITUTE FOR CLINICAL EXCELLENCE (NICE) 2004. Guide to the methods of technology appraisal. London: NICE

NATIONAL INSTITUTE FOR CLINICAL EXCELLENCE (NICE). 2007. *Antenatal and postnatal mental health: clinical management and service guidance.* [Online]. Available: www.nice.org.uk/CG45. [Accessed 6th March 2012].

NATIONAL INSTITUTE FOR CLINICAL EXCELLENCE (NICE) 2008. Guide to the methods of technology appraisal. .

NATIONAL EVALUATION OF SURE START (NESS) TEAM. 2012. The impact of sure start local programmes on seven year olds and their families. : https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/1840 73/DFE-RR220.pdf

NETTO, G., F. SOSENKO, & G. BRAMLEY. 2011. Poverty and Ethnicity in Scotland. York: Joseph Rowntree Foundation.

NEWMAN, D. A. & SIN, H. P. 2009. How Do Missing Data Bias Estimates of Within-Group Agreement? Sensitivity of SDWG, CVWG, rWG(J), rWG(J)*, and ICC to Systematic Nonresponse. *Organizational Research Methods,* 12**,** 113-147.

NEWPORT, D. J., WILCOX, M. M. & STOWE, Z. N. 2002. Maternal depression: a child's first adverse life event. *Semin Clin Neuropsychiatry,* 7**,** 113-9.

NOLEN-HOEKSEMA, S., GIRGUS, J. S. & SELIGMAN, M. E. 1992. Predictors and consequences of childhood depressive symptoms: a 5-year longitudinal study. *J Abnorm Psychol,* 101**,** 405-22.

OECD. 2011. *Doing Better for Families*. OECD Publishing.

OFFICE FOR NATIONAL STATISTICS (ONS). 2011. *Births and Deaths in England and Wales, 2011*. Statistical Bulletin. Office for National Statistics.

O'HARA, M. W. & SWAIN, A. M. 1996. Rates and risk of postpartum depression - A meta-analysis. *International Review of Psychiatry,* 8**,** 37-54.

OLDS, D. L. 2006. The Nurse–Family Partnership: An Evidence Based Preventive Intervention. . *Infant Mental Health Journal,* 27**,** 5-25.

OLDS, D. L., HENDERSON, C. R., JR., TATELBAUM, R. & CHAMBERLIN, R. 1986. Improving the delivery of prenatal care and outcomes of pregnancy: a randomized trial of nurse home visitation. *Pediatrics,* 77**,** 16-28.

OLDS, D. L., ROBINSON, J., O'BRIEN, R., LUCKEY, D. W., PETTITT, L. M., HENDERSON, C. R., JR., NG, R. K., SHEFF, K. L., KORFMACHER, J., HIATT, S. & TALMI, A. 2002. Home visiting by paraprofessionals and by nurses: a randomized, controlled trial. *Pediatrics,* 110**,** 486-96.

ORTON, J., SPITTLE, A., DOYLE, L., ANDERSON, P. & BOYD, R. 2009. Do early intervention programmes improve cognitive and motor outcomes for preterm infants after discharge? A systematic review. *Dev Med Child Neurol,* 51**,** 851-9.

PRE AND POSTNATAL DEPRESSION ADVICE AND SUPPORT (PANDAS). 2011. Postnatal depression (PND). [Online]. [Accessed 25 November 2011]. Available from http://www.pandasfoundation.org.uk

PARVIN, A., JONES, C. E. & HULL, S. A. 2004. Experiences and understandings of social and emotional distress in the postnatal period among Bangladeshi women living in Tower Hamlets. *Fam Pract,* 21**,** 254-60.

PATEL, V., RODRIGUES, M. & DESOUZA, N. 2002. Gender, poverty, and postnatal depression: a study of mothers in Goa, India. *Am J Psychiatry,* 159**,** 43-7.

PAULDEN, M., PALMER, S., HEWITT, C. & GILBODY, S. 2009. Screening for postnatal depression in primary care: cost effectiveness analysis. *BMJ,* 339**,** b5203.

PAWLBY, S., SHARP, D., HAY, D. & O'KEANE, V. 2008. Postnatal depression and child outcome at 11 years: The importance of accurate diagnosis. *Journal of Affective Disorders,* 107**,** 241-245.

PEACOCK, S., KONRAD, S., WATSON, E., NICKEL, D. & MUHAJARINE, N. 2013. Effectiveness of home visiting programs on child outcomes: a systematic review. *Bmc Public Health,* 13.

PEARLSTEIN, T. B., ZLOTNICK, C., BATTLE, C. L., STUART, S., O'HARA, M. W., PRICE, A. B., GRAUSE, M. A. & HOWARD, M. 2006. Patient choice of treatment for postpartum depression: a pilot study. *Arch Womens Ment Health,* 9**,** 303-8.

PENDRY, P. 1998. Ethological Attachment Theory: A Great Idea in Personality? [Online]. [Accessed 10 May, 2014]. Available from http://www.personalityresearch.org/papers/pendry.html.

PETROU, S. & GRAY, A. 2011. Economic evaluation alongside randomised controlled trials: design, conduct, analysis, and reporting. *BMJ,* 342**,** d1548.

PITEO, A. M., YELLAND, L. N. & MAKRIDES, M. 2012. Does maternal depression predict developmental outcome in 18 month old infants? *Early Human Development,* 88**,** 651-655.

314

PLEWIS, I., CALDERWOOD, L., HAWKES, D., HUGHES, G. & JOSHI, H. 2004. *Millennium Cohort Study First Survey: Technical Report on Sampling*. London: Centre for Longitudinal Studies.

PLEWIS, I. 2007. Non-response in a birth cohort study: the case of the Millennium Cohort Study, *International Journal of Social Research Methodology*, *10,* 325-334.

POLDERMAN, T. J., BOOMSMA, D. I., BARTELS, M., VERHULST, F. C. & HUIZINK, A. C. 2010. A systematic review of prospective studies on attention problems and academic achievement. *Acta Psychiatr Scand,* 122**,** 271-84.

POLICY, C. F. E.-B. 2010. *Nurse-Family Partnership.* [Online]. [Accessed 25th June 2011].

POOBALAN, A. S., AUCOTT, L. S., ROSS, L., SMITH, W. C., HELMS, P. J. & WILLIAMS, J. H. 2007. Effects of treating postnatal depression on mother-infant interaction and child development: systematic review. *Br J Psychiatry,* 191**,** 378-86.

POOLE, H., MASON, L. & OSBORN, T. 2006. Women's views of being screened for postnatal depression. *Community Pract,* 79**,** 363-7.

POWER, C. & ELLIOTT, J. 2006. Cohort profile: 1958 British birth cohort (National Child Development Study). *Int J Epidemiol,* 35**,** 34-41.

POWER, C. & HERTZMAN, C. 1997. Social and biological pathways linking early life and adult disease. *Br Med Bull,* 53**,** 210-21.

POWER, C., STANSFELD, S. A., MATTHEWS, S., MANOR, O. & HOPE, S. 2002. Childhood and adulthood risk factors for socio-economic differentials in psychological distress: evidence from the 1958 British birth cohort. *Social Science & Medicine,* 55**,** 1989-2004.

PRICE, J. 2008. Parent-child quality time - Does birth order matter? *Journal of Human Resources,* 43**,** 240-265.

PRIOR, V. & GLASER, D. 2006. *Understanding Attachment and Attachment Disorders: Theory, Evidence and Practice*. London and Philadelphia: Child and Adolescent Mental Health, RCPRTU.

PROUDFOOT, J., GOLDBERG, D., MANN, A., EVERITT, B., MARKS, I. & GRAY, J. A. 2003. Computerized, interactive, multimedia cognitive-behavioural program for anxiety and depression in general practice. *Psychol Med,* 33**,** 217-27.

PUGH, G. & DUFFY, B. (2013) *Contemporary Issues in the Early Years (6th Edition)*. London: Sage.

QUEVEDO, L. A., SILVA, R. A., GODOY, R., JANSEN, K., MATOS, M. B., TAVARES PINHEIRO, K. A. & PINHEIRO, R. T. 2012. The impact of maternal post-partum depression on the language development of children at 12 months. *Child Care Health Dev,* 38**,** 420-4.

RABINER, D. & COIE, J. D. 2000. Early attention problems and children's reading achievement: a longitudinal investigation. The Conduct Problems Prevention Research Group. *J Am Acad Child Adolesc Psychiatry,* 39**,** 859-67.

READING, R. & REYNOLDS, S. 2001. Debt, social disadvantage and maternal depression. *Social Science & Medicine,* 53**,** 441-453.

REDSHAW, M. & VAN DEN AKKER, O. 2007. Maternal mental health and wellbeing. *Journal of Reproductive and Infant Psychology,* 25**,** 253-254.

RENOUF, A. G. & KOVACS, M. 1994. Concordance between mothers' reports and children's self-reports of depressive symptoms: a longitudinal study. *J Am Acad Child Adolesc Psychiatry,* 33**,** 208-16.

RICHMAN, N., STEVENSON, J. & GRAHAM, P. 1982. *Preschool to school: A behavioural study. ,* London, Academic Press.

RICHTER, L., NORRIS, S., PETTIFOR, J., YACH, D. & CAMERON, N. 2007. Cohort Profile: Mandela's children: the 1990 Birth to Twenty study in South Africa. *Int J Epidemiol,* 36**,** 504-11.

RICHTERS, J. E. 1992. Depressed mothers as informants about their children: a critical review of the evidence for distortion. *Psychol Bull,* 112**,** 485-99.

RICKETTS, J., NATION, K. & BISHOP, D. V. M. 2007. Vocabulary is important for some, but not all reading skills. *Scientific Studies of Reading,* 11**,** 235-257.

ROBERTSON, E., GRACE, S., WALLINGTON, T. & STEWART, D. E. 2004. Antenatal risk factors for postpartum depression: a synthesis of recent literature. *Gen Hosp Psychiatry,* 26**,** 289-95.

RODGERS, B., PICKLES, A., POWER, C., COLLISHAW, S. & MAUGHAN, B. 1999. Validity of the Malaise Inventory in general population samples. *Social Psychiatry and Psychiatric Epidemiology,* 34**,** 333-341.

ROGOSCH, F. A., CICCHETTI, D. & TOTH, S. L. 2004. Expressed emotion in multiple subsystems of the families of toddlers with depressed mothers. *Dev Psychopathol,* 16**,** 689-709.

ROSENZWEIG, M. R. & WOLPIN, K. I. 1994. Are There Increasing Returns to the Intergenerational Production of Human-Capital - Maternal Schooling and Child Intellectual Achievement. *Journal of Human Resources,* 29**,** 670-693.

ROSS, C. E. & WU, C. 1995. The links between education and health. *American Sociological Review,* 60**,** 719-745

ROYAL COLLEGE OF MIDWIVES. 2014. *Maternal mental health - Improving emotional wellbeing in postnatal care*. Available at: https://www.rcm.org.uk/content/maternal-mental-health [Accessed 20 May 2014].

RUBIN, D.B. 1978. Multiple Imputations in Sample Surveys-A Phenomenological Bayesian Approach to Nonresponse. In *Proceedings of the Survey Research Methods Section*. American Statistical Association, pp. 20-34.

RUBEN, R. J. 1997. A time frame of critical/sensitive periods of language development. *Acta Otolaryngol,* 117**,** 202-5.

RUBIN, D. B. 1996. Multiple imputation after 18+ years. *Journal of the American Statistical Association,* 91**,** 473-489.

RUBERTSSON, C., WICKBERG, B., GUSTAVSSON, P. & RADESTAD, I. 2005. Depressive symptoms in early pregnancy, two months and one year postpartum-prevalence and psychosocial risk factors in a national Swedish sample. *Arch Womens Ment Health,* 8**,** 97-104.

RUTTER, M., TIZARD, J. & WHITMORE, K. 1970. *Education, Health and Behaviour.* London: Longmans.

RUTTER, M. & MADGE, N. 1976. *Cycles of disadvantage: a review of research. ,* London, Heinemann Educational Books.

SACKS, H. S., BERRIER, J., REITMAN, D., ANCONABERK, V. A. & CHALMERS, T. C. 1987. Meta-Analyses of Randomized Controlled Trials. *New England Journal of Medicine,* 316**,** 450-455.

SATTERFIELD, J. H. & SCHELL, A. 1997. A prospective study of hyperactive boys with conduct problems and normal boys: adolescent and adult criminality. *J Am Acad Child Adolesc Psychiatry,* 36**,** 1726-35.

SCHAFER, J. L. 1997. *Analysis of Incomplete Multivariate Data*. London: Chapman and Hall p430

SCHAFFER, H. R. 2000. The early experience assumption: Past, present, and future. . *International Journal of Behavioural Development,* 24**,** 5-14.

SCHOON, I. 2006. *Risk and resilience. Adaptations in changing times. ,* Cambridge, Cambridge University Press.

SCOTTISH INTERCOLLEGIATE GUIDELINES NETWORK (SIGN). 2012. Management of perinatal mood disorders. [Online] [Accessed 15 March 2012]. Available at http://www.sign.ac.uk/guidelines/fulltext/127

SCULPHER, M., CLAXTON, K. & AKEHURST, R. 2005. It's just evaluation for decision making: recent developments in, and challenges for, cost-effectiveness research. . *In:* SMITH, P. C., GINNELLY, L. & SCULPHER, M. (eds.) *Health policy and economics.* Maidenhead: Open University Press.

SEGRE, L. S., STASIK, S. M., O'HARA, M. W. & ARNDT, S. 2010. Listening visits: an evaluation of the effectiveness and acceptability of a home-based depression treatment. *Psychother Res,* 20**,** 712-21.

SELIGMAN, L. D., OLLENDICK, T. H., LANGLEY, A. K. & BALDACCI, H. B. 2004. The utility of measures of child and adolescent anxiety: a meta-analytic review of the Revised Children's Manifest Anxiety Scale, the State-Trait Anxiety Inventory for Children, and the Child Behavior Checklist. *J Clin Child Adolesc Psychol,* 33**,** 557-65.

SEXTON, M. B., FLYNN, H. A., LANCASTER, C., MARCUS, S. M., MCDONOUGH, S. C., VOLLING, B. L., LOPEZ, J. F., KACIROTI, N. & VAZQUEZ, D. M. 2012. Predictors of Recovery from Prenatal Depressive Symptoms from Pregnancy Through Postpartum. *Journal of Womens Health,* 21**,** 43-49.

SHAH, A. & JENKINS, R. 2000. Mental health economic studies from developing countries reviewed in the context of those from developed countries. *Acta Psychiatr Scand,* 101**,** 87-103.

SHAKESPEARE, J. 2005. *Screening: The role and recommendations of the UK National Screening Committee.,* London, Jessica Kingsley Publishers.

SHAKESPEARE, J., BLAKE, F. & GARCIA, J. 2003. A qualitative study of the acceptability of routine screening of postnatal women using the Edinburgh Postnatal Depression Scale. *Br J Gen Pract,* 53**,** 614-9.

SHEA, B. J., GRIMSHAW, J. M., WELLS, G. A., BOERS, M., ANDERSSON, N., HAMEL, C., PORTER, A. C., TUGWELL, P., MOHER, D. & BOUTER, L. M. 2007. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *Bmc Medical Research Methodology,* 7.

SHEA, B. J., HAMEL, C., WELLS, G. A., BOUTER, L. M., KRISTJANSSON, E., GRIMSHAW, J., HENRY, D. A. & BOERS, M. 2009. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. *Journal of Clinical Epidemiology,* 62**,** 1013-1020.

SHELDON, T. A. 1996. Problems of using modelling in the economic evaluation of health care. *Health Econ,* 5**,** 1-11.

SHONKOFF, J. P. & PHILLIPS, D. A. 2000. *From Neurons to Neighborhoods: The Science of Early Childhood Development* Washington, National Academies Press.

SILVA, P. A. 1990. The Dunedin Multidisciplinary Health and Development Study: a 15 year longitudinal study. *Paediatr Perinat Epidemiol,* 4**,** 76-107.

SIMONOFF, E., ELANDER, J., HOLMSHAW, J., PICKLES, A., MURRAY, R. & RUTTER, M. 2004. Predictors of antisocial personality. Continuities from childhood to adult life. *Br J Psychiatry,* 184**,** 118-27.

SINCLAIR, D. & MURRAY, L. 1998. Effects of postnatal depression on children's adjustment to school. Teacher's reports. *Br J Psychiatry,* 172**,** 58-63.

SMITH, G. D., HART, C., BLANE, D. & HOLE, D. 1998. Adverse socioeconomic conditions in childhood and cause specific adult mortality: prospective observational study. *BMJ,* 316**,** 1631-5.

SPOHR, H. L., WILLMS, J. & STEINHAUSEN, H. C. 1993. Parental alcohol exposure and long-term developmental consequences. *Lancet,* 341**,** 907-1010.

SROUFE, L. A. 1997. Psychopathology as an outcome of development. *Development and Psychopathology,* 9**,** 251-268.

SROUFE, L. A. 2005. Attachment and development: a prospective, longitudinal study from birth to adulthood. *Attach Hum Dev,* 7**,** 349-67.

STAMS, G. J., JUFFER, F. & VAN, I. M. H. 2002. Maternal sensitivity, infant attachment, and temperament in early childhood predict adjustment in middle childhood: the case of adopted children and their biologically unrelated parents. *Dev Psychol,* 38**,** 806-21.

STANTON-CHAPMAN, T. L., CHAPMAN, D. A., BAINBRIDGE, N. L. & SCOTT, K. G. 2002. Identification of early risk factors for language impairment. *Research in Developmental Disabilities,* 23**,** 390-405.

STEIN, A., MALMBERG, L. E., SYLVA, K., BARNES, J., LEACH, P. & TEAM, F. 2008. The influence of maternal depression, caregiving, and socioeconomic status in the post-natal year on children's language development. *Child Care Health and Development,* 34**,** 603-612.

STERNE, J. A., WHITE, I. R., CARLIN, J. B., SPRATT, M., ROYSTON, P., KENWARD, M. G., WOOD, A. M. & CARPENTER, J. R. 2009. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ,* 338**,** b2393.

STERNE, J. A. C. & SMITH, G. D. 2001. Sifting the evidence-what's wrong with significance tests? (Reprinted from Brit Med J, vol 322, pg 226-231, 2001). *Physical Therapy,* 81**,** 1464-1469.

STEVENSON, J. & GOODMAN, R. 2001. Association between behaviour at age 3 years and adult criminality. *Br J Psychiatry,* 179**,** 197-202.

STEWART, R. C. 2007. Maternal depression and infant growth: a review of recent evidence. *Matern Child Nutr,* 3**,** 94-107.

STOWE, Z. N. & NEMEROFF, C. B. 1995. Women at risk for postpartum-onset major depression. *Am J Obstet Gynecol,* 173**,** 639-45.

SUN, X. & FAUNCE, T. 2008. Decision-analytical modelling in health-care economic evaluations. *Eur J Health Econ,* 9**,** 313-23.

SYLVA, K. 1997. Critical periods in childhood learning. *Br Med Bull,* 53**,** 185-97.

TETI, D. M., MESSINGER, D. S., GELFAND, D. M. & ISABELLA, R. 1995. Maternal Depression and the Quality of Early Attachment - an Examination of Infants, Preschoolers, and Their Mothers. *Developmental Psychology,* 31**,** 364-376.

THOMAS, B. H., CILISKA, D., DOBBINS, M. & MICUCCI, S. 2004. A process for systematically reviewing the literature: providing the research evidence for public health nursing interventions. *Worldviews Evid Based Nurs,* 1**,** 176-84.

THOMAS, H. 1998. *Quality assessment tool for quantitative studies. Effective Public Health Practice Project. ,* Toronto, McMaster University.

THOMAS, R. & ZIMMER-GEMBECK, M. J. 2007. Behavioral outcomes of Parent-Child Interaction Therapy and Triple P-Positive Parenting Program: a review and meta-analysis. *J Abnorm Child Psychol,* 35**,** 475-95.

THOMPSON, R.A. 2006. The development of the person: Social understanding, relationships, self, conscience. In W. Damon, R.M. Lerner, & N.Eisenberg (Eds). *Handbook of Child Psychology (6th Edition)*. New York: Wiley.

UK NATIONAL SCREENING COMMITTEE. 2009. *Criteria for appraising the viability, effectiveness and appropriateness of a screening programme.* [Online]. Available: www.screening.nhs.uk/criteria [Accessed 6th March 2012].

VAN DE MHEEN, H., STRONKS, K., LOOMAN, C. W. & MACKENBACH, J. P. 1998. Does childhood socioeconomic status influence adult health through behavioural factors. *International Journal of Epidemiology,* 27**,** 431-437.

VAN DE MHEEN, H., STRONKS, K., LOOMAN, C. W. & MACKENBACH, J. P. 1998. Role of childhood health in the explanation of socioeconomic inequalities in early adult health. *J Epidemiol Community Health,* 52**,** 15-9.

VAN DE MHEEN, H., STRONKS, K., VAN DEN BOS, J. & MACKENBACH, J. P. 1997. The contribution of childhood environment to the explanation of socio-economic inequalities in health in adult life: a retrospective study. *Social Science & Medicine,* 44**,** 13-24.

VERBEEK, T., BOCKTING, C. L., VAN PAMPUS, M. G., ORMEL, J., MEIJER, J. L., HARTMAN, C. A. & BURGER, H. 2012. Postpartum depression predicts offspring mental health problems in adolescence independently of parental lifetime psychopathology. *J Affect Disord,* 136**,** 948-54.

VONDRA, J. I., SHAW, D. S., SWEARINGEN, L., COHEN, M. & OWENS, E. B. 2001. Attachment stability and emotional and behavioral regulation from infancy to preschool age. *Dev Psychopathol,* 13**,** 13-33.

WADSWORTH, M. E. J. 1991. *The imprint of time: childhood, history and adult life,* Oxford, oxford University Press.

WADSWORTH, M. E. J. 1996. Family and education as determinants of health. *In:* BLANE, D., BRUNNER, E. & WILKINSON, R. G. (eds.) *Health and social organisation.* London: Routledge.

WADSWORTH, M. E. J. 1999. Early Life. *In:* MARMOT, M. & WILKINSON, R. G. (eds.) *Social Determinants of Health.* Oxford: Oxford University Press.

WALDFOGEL, J., HAN, W. J. & BROOKS-GUNN, J. 2002. The effects of early maternal employment on child cognitive development. *Demography,* 39**,** 369-392.

WALKER, M. J., DAVIS, C., AL-SAHAB, B. & TAMIM, H. 2013. Reported maternal postpartum depression and risk of childhood psychopathology. *Matern Child Health J,* 17**,** 907-17.

WAN, M. W. & GREEN, J. 2009. The impact of maternal psychopathology on child-mother attachment. *Arch Womens Ment Health,* 12**,** 123-34.

WASHBROOK, E., PROPPER, C. & SAYAL, K. 2013. Pre-school hyperactivity/attention problems and educational outcomes in adolescence: prospective longitudinal study. *British Journal of Psychiatry,* 203**,** 265-271.

WATERS, E. & CUMMINGS, E. M. 2000. A secure base from which to explore close relationships. *Child Development,* 71**,** 164-172.

WEAVER, I. C., CERVONI, N., CHAMPAGNE, F. A., D'ALESSIO, A. C., SHARMA, S., SECKL, J. R., DYMOV, S., SZYF, M. & MEANEY, M. J. 2004. Epigenetic programming by maternal behavior. *Nat Neurosci,* 7**,** 847-54.

WEINFIELD, N. S., SROUFE, L. A. & EGELAND, B. 2000. Attachment from infancy to early adulthood in a high-risk sample: continuity, discontinuity, and their correlates. *Child Dev,* 71**,** 695-702.

WEINSTOCK, M. 2005. The potential influence of maternal stress hormones on development and mental health of the offspring. *Brain Behav Immun,* 19**,** 296-308.

WHITAKER, S. 2010. Error in the estimation of intellectual ability in the low range using the WISC-IV and WAIS-III. *Personality and Individual Differences,* 48**,** 517-521.

WHITE, F., HAYES, B., & LIVESEY, D. 2005. *Developmental Psychology: From Infancy to Adulthood*. NSW: Pearson Education Australia

WHOOLEY, M. A., AVINS, A. L., MIRANDA, J. & BROWNER, W. S. 1997. Case-finding instruments for depression. Two questions are as good as many. *J Gen Intern Med,* 12**,** 439-45.

WILSON, P., MCCONNACHIE, A. & STIRLING, M. 2002. Evolving general practice consultation in Britain - Increasing consultation time may not be straightforward. *British Medical Journal,* 325**,** 104-105.

WOODWARD, L. J. & FERGUSSON, D. M. 2000. Childhood peer relationship problems and later risks of educational under-achievement and unemployment. *J Child Psychol Psychiatry,* 41**,** 191-201.

WORLD HEALTH ORGANIZATION. 1992. *The ICD-10 classification of mental and behavioural disorders: clinical descriptions and diagnostic guidelines*. Geneva, World Health Organization.

WRIGHT, A. F. 1994. Should general practitioners be testing for depression? *Br J Gen Pract,* 44**,** 132-5.

WYNDER, E. L. 1998. Introduction to the report on the conference on the "critical" period of brain development. *Prev Med,* 27**,** 166-7.

YAWN, B. P., PACE, W., WOLLAN, P. C., BERTRAM, S., KURLAND, M., GRAHAM, D. & DIETRICH, A. 2009. Concordance of Edinburgh Postnatal Depression Scale (EPDS) and Patient Health Questionnaire (PHQ-9) to assess increased risk of depression among postpartum women. *J Am Board Fam Med,* 22**,** 483-91.

YOUTH IN MIND. 2011. Information for researchers and professionals about the Strengths & Difficulties Questionnaires. Available at: http://www.sdqinfo.org/ [Accessed 01 January 2011].

ZILL, N. 1994. Understanding Why Children in Stepfamilies Have More Learning and Behavior Problems Than Children in Nuclear Families. *Stepfamilies*, 97-106.