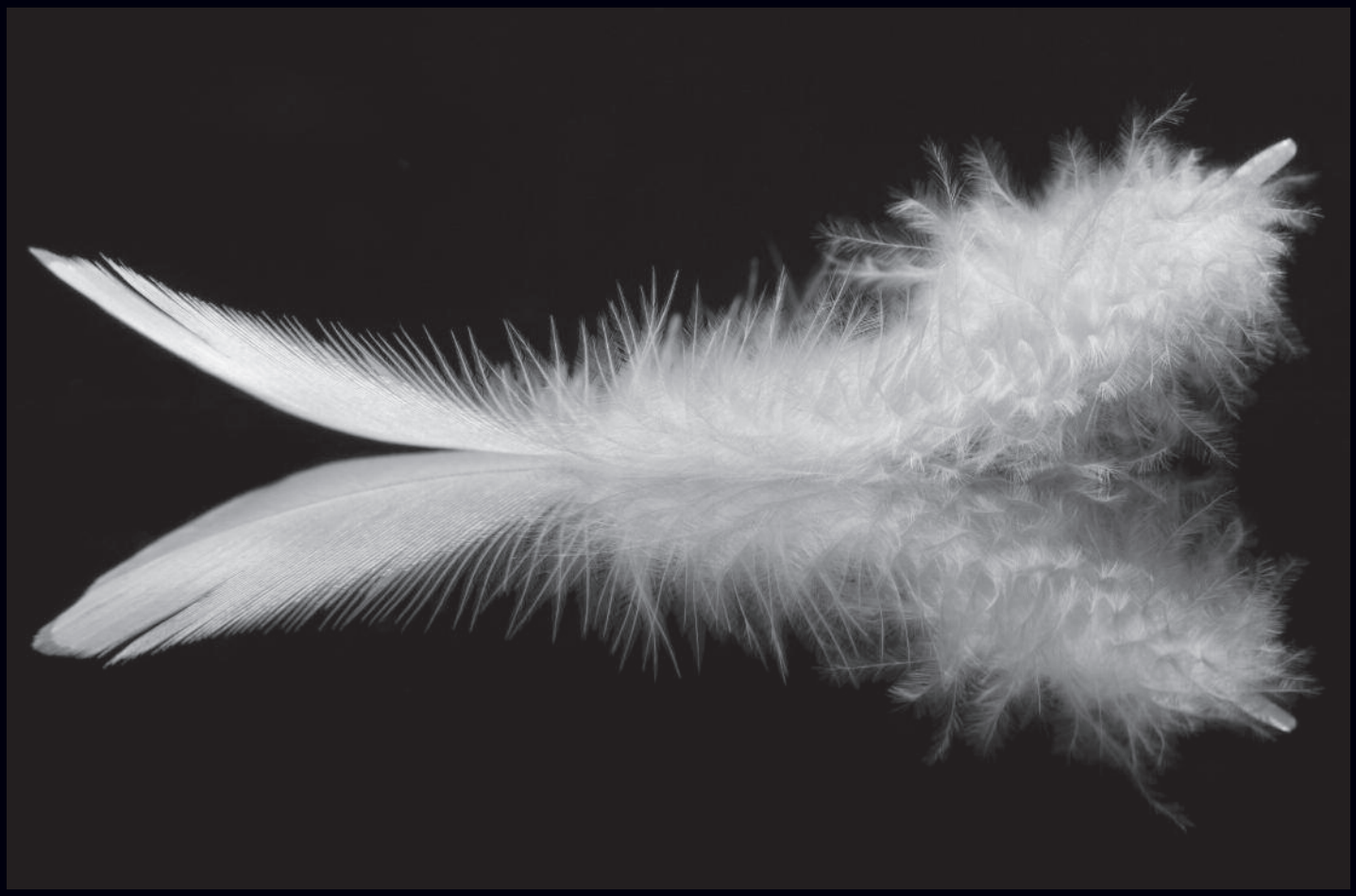


ANTENATAL AND POSTNATAL MENTAL HEALTH

THE NICE GUIDELINE ON CLINICAL MANAGEMENT AND SERVICE GUIDANCE

UPDATED EDITION

April 2018: Footnotes and cautions have been added and amended to link to the MHRA's latest advice and resources on sodium valproate. Sodium valproate must not be used in pregnancy, and only used in girls and women when there is no alternative and a pregnancy prevention plan is in place. This is because of the risk of malformations and developmental abnormalities in the baby.
See these changes in the short version of the guideline at:
http://www.nice.org.uk/guidance/CG192.

NATIONAL COLLABORATING CENTRE FOR MENTAL HEALTH

Antenatal and postnatal mental health

Clinical management and service guidance

Updated edition

This guideline should be read in conjunction with 'Service User Experience in Adult Mental Health', NICE Clinical Guidance 136 and
'Patient experience in adult NHS services', NICE Clinical Guidance 138.

National Clinical Guideline Number 192

National Collaborating Centre for Mental Health

commissioned by the

National Institute for Health and Care Excellence

published by

The British Psychological Society
and The Royal College of Psychiatrists

CONTENTS

¦ LJŘI·ŭŜ ƛyŦ2ℕYI·ŭƛ2y
! dzƎdzăŭ нлмтY C22ŭy2ŭŜă ǿŜℕŜ I·ŘŘŜŘ ŭ2 ℕŜŎ2YYŜyŘI·ŭƛ2yã мΦнΦoΣ мΦпΦнтΣ мΦпΦну I·yŘ мΦпΦнф ǿƛŭK I· fƛy⌉ ŭ2 ŭKŜ a I w ! ŭ22f⌉λŭ 2y˝ŭKŜ ℕƛă⌉ă 2Ŧ ØI·fLJℕ2I·ŭŜ YŜŘƛŎλySă ƛy˝ŦŜYI·fŜ LJI·ŭƛŜyŭăΦ ! Ŏℕ2ăăfƛy⌉ ƛy˝ ℕŜŎ2YYŜyŘI·ŭƛ2y˝мΦyΦно ǿI·ă dzLJŘI·ŭŜŘ ŭ2 fƛy⌉ ŭ2 ŭKŜ bL/9 ƎdzƛŘŜfƛySŜ 2y˝Øƛ2fŜyŎŜ I·yŘ I·ƎƎℕŜăăƛ2yΦ C22ŭy2ŭŜă ǿŜℕŜ I·ŘŘŜŘ ŭ2 ΦℕŜŎ2YYŜyŘI·ŭƛ2yã мΦпΦмт I·yŘ мΦфΦф I·ŘØƛăƛyƎ LJŜ2LJfŜ ŭKI·ŭ ŭKŜ ¦ Y 5ℕdzƎă ƛy˝ [I·ŎŭI·ŭƛ2y˝ ! ŘØƛă2ℕè ƛă I·ØI·ƛfI·ōfŜ I·ă I·y˝I·ŘŘƛŭƛ2yI·f ℕŜă2dzℕŎŜ ǿKŜy˝ăŜŜ⌉ƛyƎ I·ŘØƛŎŜ I·ō2dzŭ ăLJŜŎƛŦλŎ ŘℕdzƎăΦ

¢KŜăŜ ŎKI·yƎŜă ŎI·y˝ōŜ ăŜŜy˝ƛy˝ŭKŜ ăK2ℕŭ ØŜℕăƛ2y˝2Ŧ ŭKŜ ƎdzƛŘŜfƛySŜ I·ŭ ǿǿǿΦyƛŎŜΦ2ℕƎΦdz⌉κƎdzƛŘI·yŎŜkŎƎмфн

GUIDELINE DEVELOPMENT GROUP MEMBERS AND NATIONAL COLLABORATING CENTRE FOR MENTAL HEALTH REVIEW (NCCMH) TEAM

Antenatal and Postnatal Mental Health (Update; 2014)

Louise Howard (Chair, Guideline Development Group)
Professor in Women's Mental Health, Health Service and Population Research Department, Institute of Psychiatry, King's College London; Honorary Consultant Perinatal Psychiatrist, South London and Maudsley Foundation NHS Trust; Head of Section of Women's Mental Health and Women's Mental Health Lead for Women's Health Academic Centre, King's College London

Stephen Pilling (Facilitator, Guideline Development Group)
Director, NCCMH; Director, Centre for Outcomes Research and Effectiveness, University College London

Helen Adams
Commissioning Manager, Children and Young People Nene and Corby Clinical Commissioning Group

Jane Barlow
Professor of Public Health in the Early Years, University of Warwick

Maria Bavetta
Service User

Sonji Clarke
Senior Consultant in Gynaecology and Obstetrics, Guy's and St Thomas' Hospital, London

Asha Day
Health Visitor, East Midlands

Jill Demilew
Consultant Midwife, King's College Hospital NHS Foundation Trust, London

Karen Grayson
Service User

Alain Gregoire
Consultant Psychiatrist/Lead Clinician, Perinatal Mental Health Service, Southern Health NHS Trust; Honorary Senior Lecturer in Psychiatry, University of Southampton; Clinical Advisor for Mental Health and Learning Disability, South Central Strategic Health Authority (NHS South of England); Chair, South of England Perinatal Mental Health Network

Ian Jones
Reader in Perinatal Psychiatry and Honorary Consultant Perinatal Psychiatrist, MRC Centre for Neuropsychiatric Genetics and Genomics, Institute of Psychological Medicine and Clinical Neuroscience, Cardiff University

Liz McDonald
Consultant Perinatal Psychiatrist in East London Foundation Trust, Homerton University Hospital

Kirstie McKenzie-McHarg
Consultant Clinical Psychologist, Head of Perinatal Psychology Service, Department of Clinical Health Psychology, Warwick Hospital

Odette Megnin-Viggars
Systematic Reviewer, NCCMH (from January 2014)

Maryla Moulin
Project Manager, NCCMH

Heather O'Mahen
Senior Lecturer, School of Psychology, University of Exeter, Exeter

Sally Russell
Director and Founder of Netmums

Judith Shakespeare
Retired GP, Oxford

Eric Slade
Health Economist, NCCMH

Sarah Stockton
Senior Information Scientist, NCCMH

Iona Symington
Research Assistant, NCCMH

Clare Taylor
Senior Editor NCCMH

Amina Yesufu-Udechuku
Systematic Reviewer, NCCMH (until January 2014)

ACKNOWLEDGEMENTS

The Guideline Development Group and the NCCMH review team would like to thank the **Clinical Guidelines Technical Support Unit** and specifically the following people:

Ron Gray
Senior Clinical Research Fellow and Consultant Clinical Epidemiologist, National Perinatal Epidemiology Unit, University of Oxford; Honorary Consultant in Public Health, Oxfordshire Primary Care Trust

Anne Lingford-Hughes
Professor of Addiction Biology at Imperial College; Consultant Psychiatrist with a particular interest in pharmacological treatments of alcohol problems and comorbidity at Central North West London NHS Foundation Trust

Nadia Micali
NIHR Clinician Scientist (Senior Lecturer), Behavioural and Brain Sciences Unit, Institute of Child Health, University College London, London; Honorary Consultant Psychiatrist, Great Ormond Street Hospital, London

Anne Oxley
Team Manager, Perinatal Mental Health Services, Northumberland, Tyne and Wear NHS Trust

Assistance in reviewing the health economic literature

Toshi Baba
Honorary Research Associate, University College London

Assistance in reviewing the clinical evidence

Rebecca Gate
Research Assistant, NCCMH

Bronwyn Harrison
Systematic Reviewer, NCCMH

Elena Marcus
Research Assistant, NCCMH

Cheryl Palmer
Research Assistant, NCCMH

Editorial assistance

Nuala Ernest
Assistant Editor, NCCMH

GUIDELINE DEVELOPMENT GROUP MEMBERS 2007

Antenatal and Postnatal Mental Health (2007)

Dave Tomson (Chair)
GP and Consultant in Patient-Centred Primary Care, North Shields

Stephen Pilling (Facilitator)
Joint Director, NCCMH; Director, Centre for Outcomes Research and Effectiveness, University College London; Consultant Clinical Psychologist, Camden and Islington Mental Health and Social Care Trust

Fiona Blake
Consultant Psychiatrist, Cambridge University Hospitals NHS Foundation Trust

Rachel Burbeck
Systematic Reviewer, NCCMH (from July 2005)

Sandra Elliott
Consultant Clinical Psychologist, South London and Maudsley NHS Trust

Pauline Evans
Service User Representative

Josephine Foggo
Project Manager, NCCMH (until August 2005)

Alain Gregoire
Consultant Perinatal Psychiatrist, Hampshire Partnership NHS Trust and University

Jane Hamilton
Consultant Psychiatrist in Maternal Health, Sheffield Care Trust

Claire Hesketh
Primary Care Mental Health Manager, Northumberland, Tyne and Wear NHS Trust

Rebecca King
Project Manager, NCCMH (from August 2005)

Elizabeth McDonald
Consultant Perinatal Psychiatrist, East London and the City NHS Mental Health Trust

Rosa Matthews
Systematic Reviewer, NCCMH (until July 2005)

Ifigeneia Mavranezouli
Health Economist, NCCMH

Patrick O'Brien
Obstetrician, University College London Hospitals NHS Foundation Trust

Donald Peebles
Obstetrician, University College London Hospitals NHS Foundation Trust

Sue Power
Team Manager for Community Mental Health Team, Vale of Glamorgan County Council

Yana Richens
Consultant Midwife, University College London Hospitals NHS Foundation Trust

Ruth Rothman
Specialist Health Visitor for Postnatal Depression and Clinical Lead for Mental Health, Southend Primary Care Trust

Fiona Shaw
Service User Representative

Sarah Stockton
Information Scientist, NCCMH

Clare Taylor
Senior Editor, NCCMH

Lois Thomas
Research Assistant, NCCMH (until September 2005)

Jenny Turner
Research Assistant, NCCMH (from November 2005)

1 PREFACE

This guideline was first published in February 2007. This edition of the guideline updates most areas, except for the organisation of services, although the vignettes within the chapter on organisation of services (Chapter 4) have been removed because a new review of the experience of care has been conducted (see Chapter 6). The chapter entitled 'Prediction and detection of mental illnesses during pregnancy and the postnatal period' from the 2007 guideline has also been removed because prediction was not in the remit for this guideline and detection was covered in Chapter 5.

This guideline has been developed to advise on the clinical management of and service provision for mental health problems in pregnancy and the postnatal period. The guideline recommendations have been developed by a multidisciplinary team of healthcare professionals, women who have experienced a mental health problem in pregnancy or the postnatal period, and the guideline methodologists, after careful consideration of the best available evidence. It is intended that the guideline will be useful to clinicians and service commissioners in providing and planning high-quality care for women with a mental health problem in pregnancy or the postnatal period while also emphasising the importance of improving the experience of care of women and their partners, families or carers (see Appendix 1 for more details on the scope of the guideline).

Although the evidence base is rapidly expanding, there are some major gaps. The guideline makes a number of research recommendations specifically to address these gaps. In the meantime, it is hoped that the guideline will assist clinicians, and women with a mental health problem in pregnancy or the postnatal period and their partners, families or carers, by identifying the merits of particular treatment approaches where the evidence from research and clinical experience exists.

1.1 NATIONAL CLINICAL GUIDELINES

1.1.1 What are clinical guidelines?

Clinical guidelines are 'systematically developed statements that assist clinicians and service users in making decisions about appropriate treatment for specific conditions' (Mann, 1996). They are derived from the best available research evidence, using predetermined and systematic methods to identify and evaluate the evidence relating to the specific condition in question. Where evidence is lacking, the guidelines include statements and recommendations based upon the consensus statements developed by the Guideline Development Group (GDG).

Clinical guidelines are intended to improve the process and outcomes of healthcare in a number of different ways. They can:

- provide up-to-date evidence-based recommendations for the management of conditions and disorders by healthcare professionals
- be used as the basis to set standards to assess the practice of healthcare professionals
- form the basis for education and training of healthcare professionals
- assist service users and their carers in making informed decisions about their treatment and care
- improve communication between healthcare professionals, service users and their carers
- help identify priority areas for further research.

1.1.2 Uses and limitations of clinical guidelines

Guidelines are not a substitute for professional knowledge and clinical judgement. They can be limited in their usefulness and applicability by a number of different factors: the availability of high-quality research evidence, the quality of the methodology used in the development of the guideline, the generalisability of research findings and the uniqueness of individuals.

Although the quality of research in this field is variable, the methodology used here reflects current international understanding on the appropriate practice for guideline development (Appraisal of Guidelines for Research and Evaluation Instrument [AGREE]; www.agreetrust.org; AGREE Collaboration, 2003), ensuring the collection and selection of the best research evidence available and the systematic generation of treatment recommendations applicable to the majority of women with a mental health problem in pregnancy or the postnatal period. However, there will always be some people and situations where clinical guideline recommendations are not readily applicable. This guideline does not, therefore, override the individual responsibility of healthcare professionals to make appropriate decisions, in consultation with the women and, if she agrees, her partner, family or carer.

In addition to the clinical evidence, cost-effectiveness information, where available, is taken into account in the generation of statements and recommendations in clinical guidelines. While national guidelines are concerned with clinical and cost effectiveness, issues of affordability and implementation costs are to be determined by the National Health Service (NHS).

In using guidelines, it is important to remember that the absence of empirical evidence for the effectiveness of a particular intervention is not the same as evidence for ineffectiveness. In addition, and of particular relevance in mental health, evidence-based treatments are often delivered within the context of an overall treatment programme including a range of activities, the purpose of which may be to help engage the person and provide an appropriate context for the delivery of specific interventions. It is important to maintain and enhance the service context in which these interventions are delivered, otherwise the specific benefits of effective interventions will be lost. Indeed, the importance of organising care in order to

support and encourage a good therapeutic relationship is at times as important as the specific treatments offered.

1.1.3 Why develop national guidelines?

The National Institute for Health and Care Excellence (NICE) was established as a Special Health Authority for England and Wales in 1999, with a remit to provide a single source of authoritative and reliable guidance for service users, professionals and the public. NICE guidance aims to improve standards of care, diminish unacceptable variations in the provision and quality of care across the NHS, and ensure that the health service is person-centred. All guidance is developed in a transparent and collaborative manner, using the best available evidence and involving all relevant stakeholders.

NICE generates guidance in a number of different ways, three of which are relevant here. First, national guidance is produced by the Technology Appraisal Committee to give robust advice about a particular treatment, intervention, procedure or other health technology. Second, NICE commissions public health intervention guidance focused on types of activity (interventions) that help to reduce people's risk of developing a disease or condition, or help to promote or maintain a healthy lifestyle. Third, NICE commissions the production of national clinical guidelines focused upon the overall treatment and management of a specific condition. To enable this latter development, NICE has established four National Collaborating Centres in conjunction with a range of professional organisations involved in healthcare.

1.1.4 From national clinical guidelines to local protocols

Once a national guideline has been published and disseminated, local healthcare groups will be expected to produce a plan and identify resources for implementation, along with appropriate timetables. Subsequently, a multidisciplinary group involving commissioners of healthcare, primary care and specialist mental health professionals, service users and carers should undertake the translation of the implementation plan into local protocols, taking into account both the recommendations set out in this guideline and the priorities in the *National Service Framework for Mental Health* (Department of Health, 1999a) and related documentation. The nature and pace of the local plan will reflect local healthcare needs and the nature of existing services; full implementation may take a considerable time, especially where substantial training needs are identified.

1.1.5 Auditing the implementation of clinical guidelines

This guideline identifies key areas of clinical practice and service delivery for local and national audit. Although the generation of audit standards is an important and necessary step in the implementation of this guidance, a more broadly-based implementation strategy will be developed. Nevertheless, it should be noted that the Care Quality Commission in England, and the Healthcare Inspectorate Wales, will monitor the extent to which commissioners and providers of health and social care and Health Authorities have implemented these guidelines.

1.2 THE NATIONAL ANTENATAL AND POSTNATAL MENTAL HEALTH GUIDELINE

1.2.1 Who has developed this guideline?

This guideline has been commissioned by NICE and developed within the National Collaborating Centre for Mental Health (NCCMH). The NCCMH is a collaboration of the professional organisations involved in the field of mental health, national service user and carer organisations, a number of academic institutions and NICE. The NCCMH is funded by NICE and is led by a partnership between the Royal College of Psychiatrists and the British Psychological Society's Centre for Outcomes Research and Effectiveness, based at University College London.

The GDG was convened by the NCCMH and supported by funding from NICE. The GDG included women who have experienced a mental health problem in the pregnancy or the postnatal period, and professionals from psychiatry, clinical psychology, general practice, nursing, health visitors, obstetrics, midwifery and the private and voluntary sectors, and a mother infant specialist.

Staff from the NCCMH provided leadership and support throughout the process of guideline development, undertaking systematic searches, information retrieval, appraisal and systematic review of the evidence. Members of the GDG received training in the process of guideline development from NCCMH staff, and the service users and carers received training and support from the NICE Patient and Public Involvement Programme. The NICE Guidelines Technical Adviser provided advice and assistance regarding aspects of the guideline development process.

All GDG members made formal declarations of interest at the outset, which were updated at every GDG meeting. The GDG met a total of twelve times throughout the process of guideline development. It met as a whole, but key topics were led by a national expert in the relevant topic. The GDG was supported by the NCCMH technical team, with additional expert advice from special advisers where needed. The group oversaw the production and synthesis of research evidence before presentation. All statements and recommendations in this guideline have been generated and agreed by the whole GDG.

1.2.2 For whom is this guideline intended?

This guideline will be relevant for women with a mental health problem in pregnancy or the postnatal period and covers the care provided by primary, community, secondary, tertiary and other healthcare professionals who have direct contact with, and make decisions concerning the care of, women with a mental health problem in pregnancy or the postnatal period.

In summary, the guideline is intended for use by:

- Professional groups who share in the treatment and care for women with a mental health problem in pregnancy or the postnatal period, including psychiatrists, clinical psychologists, mental health nurses, community psychiatric nurses (CPNs), other community nurses, general practitioners (GPs), midwives, neonatologists, obstetricians, health visitors, social workers, counsellors, practice nurses, occupational therapists, pharmacists and others.
- Professionals in other health and non-health sectors who may have direct contact with or are involved in the provision of health and other public services for women with a mental health problem in pregnancy or the postnatal period; these may include accident and emergency staff, paramedical staff, prison doctors, the police and professionals who work in the criminal justice and education sectors.
- Those with responsibility for planning services for women with a mental health problem in pregnancy or the postnatal period, and their partners, families or carers, including directors of public health and NHS Trust managers.

1.2.3 Specific aims of this guideline

The guideline makes recommendations for the use of pharmacological, psychological and service-level interventions. It aims to:

- evaluate the role of specific pharmacological agents in the treatment and management mental health problems in pregnancy and the postnatal period
- evaluate the role of specific psychological interventions in the treatment and management of mental health problems in pregnancy and the postnatal period
- evaluate the role of specific service-delivery systems and service-level interventions in the management of mental health problems in pregnancy and the postnatal period
- to provide best-practice advice on the care of women with a mental health problem in pregnancy or the postnatal period through the different phases of illness, including the initiation of treatment, the treatment of acute episodes and the promotion of recovery
- consider economic aspects of various standard treatments of mental health problems in pregnancy and the postnatal period
- promote the implementation of best clinical practice through the development of recommendations tailored to the requirements of the NHS in England and Wales.

1.2.4 The structure of this guideline

The guideline is divided into chapters, each covering a set of related topics. The first three chapters provide a general introduction to guidelines, an introduction to the

topic of mental health problems in pregnancy and the postnatal period, and to the methods used to develop this guideline. Chapters 4 to 8 provide the evidence that underpins the recommendations about the experience of care and the treatment and management of mental health problems in pregnancy and the postnatal period.

Each evidence chapter begins with a general introduction to the topic that sets the recommendations in context. Depending on the nature of the evidence, narrative reviews or meta-analyses were conducted, and the structure of the chapters varies accordingly. Where appropriate, details about current practice, the evidence base and any research limitations are provided. Where meta-analyses were conducted, information is given about both the interventions included and the studies considered for review. Clinical summaries are then used to summarise the evidence presented. Finally, recommendations related to each topic are presented at the end of each chapter or after each evidence review within a chapter. Full details about the included studies can be found in Appendix 18. Where meta-analyses were conducted, the data are presented using forest plots in Appendix 19. GRADE evidence profiles can be found in Appendix 22 and evidence tables for economic studies in Appendix 20 and Appendix 21.

Table 1: Clinical and economic evidence appendices

Evidence tables for economic studies	Appendix 20, 21
Clinical study characteristics tables	Appendix 17, 18
Clinical evidence forest plots	Appendix 19
GRADE evidence profiles	Appendix 22

In the event that amendments or minor updates need to be made to the guideline, please check the NCCMH website (nccmh.org.uk), where these will be listed and a corrected PDF file available to download.

2 ANTENATAL AND POSTNATAL MENTAL HEALTH

2.1 SCOPE OF THE GUIDELINE

This guideline covers the mental health care of women who have, or are at risk of, mental health problems in the perinatal period, which comprises pregnancy (the 'antenatal period') and the 'postnatal period' (from childbirth to the end of the first postnatal year) – the period that defines most specialist perinatal mental health services.

The guideline is concerned with a broad range of mental health problems, including depression, anxiety disorders, eating disorders, drug and alcohol-use disorders and severe mental illness (such as psychosis, bipolar disorder, schizophrenia and severe depression). This includes women with subthreshold symptoms and those with mild, moderate and severe mental health problems. However, the guideline focuses on the aspects of their expression, risks and management that are of special relevance in pregnancy and the postnatal period. Thus, the guidelines should be used in conjunction with other NICE guidance about specific mental health problems (see www.nice.org.uk/guidance/index.jsp?action=bytopic&o=7281).

The guideline also makes recommendations about the services required to support the delivery of effective identification and treatment of most mental health problems in pregnancy and the postnatal period in primary and secondary care. It will also be relevant to (but not make specific recommendations for) non-NHS services such as social services and the independent sector.

The optimisation of psychological wellbeing, as opposed to the management of mental health problems, is not covered in this guideline, however, the importance of this is implicit. The mental health needs of fathers, partners, other carers and children, whose health and functioning will inevitably be affected by mental health problems in women, are also important and should not be neglected, and their needs have been considered in developing the recommendations in this guideline. In relevant places, the phrase 'partner, family or carer' has been used to remind readers of the continued importance of thinking about mental health problems and their impact on the family.

The context of care, namely pregnancy and the postnatal period, is the primary focus of the guideline, rather than significant differences in the nature of particular mental health problems during these periods. The biological, physiological, psychological and social changes that occur at this time influence the nature of both the identification and treatment of mental health problems. Much of the guideline is concerned with the balancing of the risks and benefits of treatment at a particularly critical time in the lives of women, the fetus/baby, and their families.

2.2 MENTAL HEALTH PROBLEMS IN PREGNANCY AND THE POSTNATAL PERIOD

2.2.1 Introduction

Pregnancy and the period from childbirth to the end of the first postnatal year comprise one of the most important times of a woman's life, but for women with a mental health problem it can be difficult and distressing. In pregnancy and the postnatal period, women are vulnerable to having or developing the same range of mental health problems as other women, and the nature and course of the large majority of these problems are similar in women at other times of their lives. However, the nature and treatment of mental health problems in pregnancy and the postnatal period differ in a number of important respects:

- Women might not want to tell anyone about their feelings because of the stigma of mental health problems during a period that is broadly associated with happiness; they might also worry that social care will become involved, which they might fear could lead to loss of custody (Dolman et al., 2013).
- There is a risk of pregnant women with an existing mental health problem stopping medication, often abruptly and without the benefit of an informed discussion, which can precipitate or worsen an episode.
- In women with an existing mental health problem (for example, bipolar disorder), there is an increased risk of developing an episode during the early postnatal period. There are also some other differences in epidemiology, which are reviewed for the specific disorders below.
- The impact of any mental health problem may often require more urgent intervention than would usually be the case because of its potential effect on the fetus/baby and on the woman's physical health and care, and her ability to function and care for her family.
- Postnatal-onset psychotic disorders may have a more rapid onset with more severe symptoms than psychoses occurring at other times (Wisner & Wheeler, 1994) and demand an urgent response.
- The effects of mental health problems at this time require that not only the needs of the woman but also those of the fetus/baby, siblings and other family members are considered (including the physical needs of the woman or fetus/baby) – for example, when considering waiting times for psychological interventions, acute treatment for severe mental illnesses or admission to an inpatient bed.
- The shifting risk-benefit ratio in the use of psychotropic medication during pregnancy and the postnatal period (particularly when breastfeeding) requires review of the thresholds for treatment for both pharmacological and psychological interventions. This may result in a greater prioritisation of prompt and effective psychological interventions.

2.2.2 Course and prognosis of mental health problems in the pregnancy and the postnatal period

There is little evidence that the underlying course of most pre-existing mental health problems is significantly altered during this time, with the exception of bipolar disorder (which shows an increased rate of relapse and first presentation, see Section 2.3.4), and lower rates for alcohol-use disorders (Vesga-Lopez et al., 2008). There is also some emerging evidence to suggest that the prevalence of adjustment disorder and generalised anxiety disorder (GAD) may be higher in pregnancy and the postnatal period (Ross & MacLean, 2006) and incident depression higher in the postnatal period (Ban et al., 2012; Munk-Olsen et al., 2006). Similarly, there is little evidence that the prognosis of mental health problems that develop in pregnancy or postnatally are significantly different from those developing at other times (Brockington, 1996). However, there is evidence of increased risk of adverse outcomes for the fetus/baby, and subsequently in childhood (see Chapter 5) and an increased risk of mental health problems in the partners of women with mental health problems in pregnancy and the postnatal period (Lovestone & Kumar, 1993).

The concept of prognosis must therefore be extended to consideration of not only the future course of the mental health problem and its impact on the woman, but also its impact on the other family members. The increased vulnerability of children whose parents have a mental health problem (Beardslee et al., 1983; Rubovits, 1996; Gray, 2013) argues strongly for the effective and prompt treatment of mental health problems in pregnancy and the postnatal period. There are many opportunities for pregnant or postnatal women to be identified and treated because they are in frequent contact with universal services (maternity, health visiting and primary care) for their and their baby's care. However, healthcare professionals should also consider that women with a mental health problem may be less likely to access regular physical care, and for those who do, many might have considerable anxiety about disclosing a mental health problem. The focus on the needs of the fetus/baby by both the mother and healthcare professionals should not obscure the needs of the mother.

2.2.3 Pregnancy and birth in England and Wales

There were 729,674 live births in England and Wales in 2012 (812,970 in the UK). Over the last 10 years fertility levels have risen for women in all age groups with the exception of those aged under 20, and the total fertility rate is now 1.94 children per woman. The percentage of live births in England and Wales born to mothers born outside the UK is 25.9% compared with 11.6% in 1990. In 2012, the average age of women giving birth was 29.8, with average age for first births 28.1; 84% of babies were registered by parents who were married, in a civil partnership or cohabiting (based on figures provided by the Office for National Statistics, 2012a).

Sociodemographic factors impact on maternal and infant morbidity and mortality. In the period 2006-08 there were 0.067 maternal deaths per 1000 live births (compared with 0.13 maternal deaths per 1000 live births in 2000); women with unemployed

husbands or partners are six times more likely to die than those whose husbands or partners are employed.

In 2011 infant mortality was at its lowest ever rate (4.1 deaths per 1000 live births; Office for National Statistics, 2012b), but rates were higher (5.4 deaths per 1000 live births) among babies of mothers aged under 20 and over 40 years. Prematurity is also related to young and old maternal age, and other risk factors include socioeconomic status and educational level, ethnicity and single marital status (Goldenberg et al., 2008). The stillbirth rate in 2011 was 4.9 per 1000 deliveries but stillbirth rates are twice as high in the most deprived tenth of women compared with the least deprived tenth (Seaton et al., 2012).

In 2011, according to figures from the Office for National Statistics (2013), 7.1% of births were preterm (under 37 weeks' gestation) and of these, 1.3% were born before 24 weeks. The majority (95%) occur after 28 weeks. Nearly 5% of all babies born prematurely will have a very low birthweight (less than 1000 g), compared with 93.7% born under 24 weeks. Fewer than 1% of babies born at full term will be of very low birthweight. Young maternal age and deprivation are associated with prematurity (Taylor-Robinson et al., 2011).

Sociodemographic factors, therefore, are distal determinants of adverse pregnancy outcomes and also play an important role in both the aetiology and maintenance of mental health problems. The above figures serve to emphasise the vulnerability of some women and their babies. Such adversity may also play an important role in the maintenance of mental health problems in adults (Skapinakis et al., 2006).

2.2.4 Consequences of mental health problems in pregnancy and the postnatal period

Consequences for the woman

For a woman who develops a mental health problem, either in pregnancy or the postnatal period, there are concerns and difficulties for her in addition to those arising specifically from the mental health problem. Women can be concerned that the mental health problem may have a negative impact on the wellbeing of their fetus/baby. This can exacerbate an already disabling mental health problem. Mental health problems, particularly in their more severe form, can also be associated with significant impairment in social and personal functioning, which might have a detrimental effect on the woman's ability to care effectively for herself and her children. The impact of this can be seen most obviously and tragically in the significant number of women with schizophrenia who lose custody of their children (Howard, 2005). The long-term effects of this on the woman are considerable. Psychiatric causes of maternal death, particularly suicide, continue to be a significant cause of maternal mortality in the UK as revealed by the *Confidential Enquiries into Maternal Deaths in the UK* (Oates & Cantwell, 2011). More rarely, severe mental illness, particularly in the first postnatal month, may lead to infanticide (Flynn et al., 2007).

Consequences for the pregnancy and baby

All pregnancies carry risk, in particular to the fetus. According to statistics from Springett and colleagues (2013), there was a birth prevalence of congenital malformations of 219 per 10,000 total births (one in 46 total births) in England and Wales in 2011. Congenital anomalies contribute to an estimated 15% of infant mortality, particularly congenital heart defects (47%), chromosomal anomalies (19%) and digestive system anomalies (17%). Mothers between 25 and 29 years of age had the lowest birth prevalence for all anomalies. The prevalence was higher in the under 20 age group and considerably higher in the 40 and over age group. As discussed above, stillbirths account for 4.9 of every 1000 deliveries and around 7% are preterm.

These risks may increase if the woman has a mental health problem. There is evidence that mental health problems in pregnancy and the postnatal period are associated with adverse outcomes for the fetus and the baby as well as for the woman herself. For example, severe depression is associated with an increased risk of lower birthweight and premature babies, particularly in settings of socioeconomic deprivation (Grote et al., 2010), self-harm and suicide (Lindahl et al., 2005). In schizophrenia and bipolar disorder, there is also a risk of poorer obstetric outcomes, including placental abnormalities (abruption of the placenta, placenta previa) disorder), increased preterm delivery, low-birthweight babies and babies who are small for gestational age (Howard, 2005; Jablensky et al., 2005), increased risk of stillbirth (Webb et al., 2005; King-Hele et al., 2009) and neonatal mortality (Howard, 2005; King-Hele et al., 2009), potentially significant exacerbation of the disorder if not treated, and suicide (Oates & Cantwell, 2011). Similarly, low birthweight has been associated with maternal history of anorexia nervosa (Solmi et al., 2014) and women with binge eating disorder have an elevated risk of babies that are large for gestational age (Bulik et al., 2009). Elevated risks of sudden infant death syndrome have also been reported in relation to depression in pregnancy (Howard et al., 2007) and the postnatal period (Mitchell et al., 1992; Sanderson et al., 2002) and to maternal schizophrenia (Bennedsen et al., 2001). As with other adverse outcomes, there does not appear to be diagnostic specificity, although worse fetal and infant outcomes are often reported for drug and alcohol-use disorders (for example King-Hele et al., 2007; King-Hele et al., 2009).

There is also emerging evidence that untreated mental health problems in pregnancy may be associated with poorer long-term outcomes for children beyond the immediate postnatal period (Nulman et al., 2002). For example, depression in pregnancy has been associated with internalising and externalising disorders in the children (Barker et al., 2011; Laurent et al., 2013), and depression in adolescents and young adults (Pawlby et al., 2009; Pearson et al., 2013a); and anxiety in pregnancy is associated with an increased risk of internalising problems (Barker et al., 2011; Blair et al., 2011), and emotional and behavioural difficulties in children (O'Connor et al., 2002; O'Connor et al, 2003a).

Postnatal mental health problems in women, if chronic, can be associated with adverse cognitive outcomes for their children and mental health problems (Sutter et al., 2011) (see Chapter 5). One of the key mediating mechanisms for adverse developmental outcomes in the child appears to be impaired mother-infant interactions (Field, 2010). Severe mental illness, such as maternal schizophrenia are also associated with significant parenting difficulties (Wan et al., 2008), with a high proportion of women losing care of their baby (Howard et al., 2004)

Although there is an increased risk of adverse outcomes in the children of mothers with mental health problems, these are not inevitable. It is difficult to establish whether many of the associations are causal because large sample sizes are needed to disentangle the effect of mental health problems in pregnancy and the postnatal period from other risk factors. There is growing evidence, for example, that socioeconomic adversity, socioeconomic status and education modify the association between depression in the postnatal period and child outcomes; that is, poor outcomes occur predominantly in families living in socioeconomic difficulties (Pearson et al., 2013a; Lovejoy et al., 2000). Recent research has reported that personality disorder may moderate the impact of mental health problems on child outcomes – dysregulated infant behaviour occurs in children of women with depression who have a personality disorder, but not in children of women with depression but no personality disorder (Conroy et al., 2012). It is also possible that risk factors such as smoking, obesity or domestic violence and abuse, which are more common in women with mental health problems, explain some of the adverse consequences of mental health problems in pregnancy and the postnatal period because these comorbidities are also risk factors for adverse child outcomes.

Coupled with the direct effects of maternal mental health problems on the fetus and baby, there are important indirect effects such as social isolation and other disadvantages known to be associated with severe mental illness, in addition to genetic risk of mental health problems. All of these factors point to the importance of appropriate and timely treatment of the woman during pregnancy, and the woman and the baby in the postnatal period.

Both psychological and pharmacological interventions are effective in the treatment of most major mental health problems (NICE 2004a, NICE 2005a, NICE 2009, NICE 2011a, NICE 2014). For a proportion of women, where psychological treatment alone may be insufficient and medication is needed as prophylaxis or treatment, pharmacological interventions may be the treatment both advocated by a healthcare professional and chosen by the woman herself. The evidence for the possible risk from different medications to the baby is reviewed in Chapter 8. However, as has been described above, untreated mental health problems may also impact adversely on the fetus/baby. For women and clinicians, the assessment of drug treatment risk is therefore highly complex and further complicated by the need to balance this against the harm of the untreated mental health problem. In addition to possible teratogenic and other risks to the fetus, such as smoking or alcohol use, the altered physical state of the woman over the course of a pregnancy means that increased

physical monitoring, for example drug levels for medications that will change during the course of pregnancy, and the impact on breastfeeding, all need to be considered when making decisions about pharmacological treatment. These issues are discussed more fully in Chapter 8.

2.3 INCIDENCE AND PREVALENCE OF MENTAL HEALTH PROBLEMS IN PREGNANCY AND THE POSTNATAL PERIOD

The purpose of this section is not to provide an exhaustive overview of the epidemiology of mental health problems in pregnancy and the postnatal period but to highlight important issues about their incidence and prevalence, particularly if they are different from those found in general adult populations. The commentary below is also limited as a result of the paucity of research in this area. Most studies to date have focused principally on depression and psychotic disorders, mainly in the postnatal period, and studies of depression have generally relied on the use of self-report measures applied at isolated time points. Therefore, caution must be applied to the interpretation of the data and to the use of the term 'postnatal depression' (or 'postpartum depression'). There is concern that this term is used in clinical situations as a label for any mental health problem occurring in the postnatal period and as a consequence other severe mental illnesses might fail to be identified (Lewis & Drife, 2004). It also reinforces the view that depression in the postnatal period is somehow distinct from depression at other times. Common false beliefs about depression in the postnatal period include the idea that its symptoms and effects are always less severe, that it usually goes away by itself, that it is caused by hormone levels, that it has no risk of non-postnatal recurrence, that it carries an inevitable risk of future postnatal recurrence, or that it is different from depression that is already present before childbirth. All of these assumptions are misleading and can lead to disadvantageous and inappropriate responses by clinicians and women themselves. In addition, they can lead to policy and service development focused on depression postnatally, to the exclusion of the full range of mental health problems occurring in pregnancy and the postnatal period, all of which can potentially have serious effects on the woman, her fetus/baby and her family.

It is therefore recommended that, for the purpose of diagnosis, usual diagnostic guidelines for each condition, such as those contained in *The ICD-10 Classification of Mental and Behavioural Disorders* (ICD-10) (World Health Organization [WHO], 1992) and the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-V) (American Psychiatric Association, 2013), be followed. Clinicians should bear in mind that some changes in mental state and functioning are a normal part of pregnancy and the postnatal experience and should, therefore, be cautious about basing any diagnosis largely on such features without careful consideration of the context. Such features include appetite change, which is a poor indicator of depression in pregnancy and the postnatal period (Kammerer et al., 2009; Nylen et al., 2013); but sleep disturbance, tiredness, loss of libido and anxious thoughts about

the baby may also be considered 'normal' whereas careful clinical assessment may reveal a mental health problem.

2.3.1 Depression

Depression is common and is associated with major disability when following a chronic course (WHO, 1992), but it is not the only mental health problem in pregnancy or the postnatal period, despite its dominance in the perinatal mental health literature. The estimated point prevalence for major depression among 16- to 65-year-olds in the UK is 21 per 1000 (males 17, females 25); but, if the less specific and broader category of 'mixed depression and anxiety' (F41.2, ICD-10, WHO, 1992) is included, these figures rise dramatically to 98 per 1000 (males 71, females 124). In mixed depression and anxiety, it can be seen that the gender ratio is more skewed to females (Meltzer et al., 1995a; Meltzer et al., 1995b). Differential rates of prevalence of depression are identified in the same study, being highest among the separated (56 per 1000 female, 111 per 1000 male), next highest among widowed males (70 per 1000) and divorced females (46 per 1000), with the lowest prevalence among the married (17 and 14 per 1000, respectively). Lone parents have higher rates than couples, and couples with children higher rates than those without children (Meltzer et al., 1995a; Meltzer et al., 1995b). Socioeconomic deprivation is associated with depression, with recent research indicating that this is also found for depression in pregnancy and the postnatal period (Ban et al., 2012). Epidemiological studies have also established that, for most, depression is chronic. In a WHO study, 66% of those identified as having depression were still found to satisfy criteria for a mental health problem 1 year later, and for 50% the diagnosis was depression. It is probable that widely differing rates between the clinics studied in the countries in which the data were collected reflect true differences in prevalence in these clinics rather than differing concepts of depression between countries (Simon et al., 2002a).

Although research and clinical care has generally placed the greatest emphasis on the postnatal period, depression in pregnancy is also of considerable importance. A high-quality review of depression in pregnancy and the postnatal period, which used meta-analysis to combine point prevalence estimates from large-scale studies, estimated the point prevalence of major depression (that is, the rate at a particular point in time) as 3.8% at the end of the first trimester, 4.9% at the end of the second and 3.1% at the end of the third (Gavin et al., 2005). The same review estimated the postnatal point prevalence at between 1 and 5.7% in the first postnatal year, with the highest rates at 2 months (5.7%) and 6 months (5.6%) postnatally. Gavin and colleagues calculated the period prevalence (that is, the rate over a period of time) as 12.7% in pregnancy, 5.7% from birth to 2 months postnatally, 6.5% at 6 months and 21.9% at 12 months. However, for most of these estimates, only a single study was found. The estimates contrast with a large-scale community prospective study of around 8,300 women (based on the *Avon Longitudinal Study of Parents and Children* [ALSPAC; O'Connor et al., 2003b; Heron et al., 2004]), which measured depressive symptoms in pregnancy and the postnatal period (from 18 weeks' gestation to 8 months postnatally), and found that depression scores were higher at 32 weeks' gestation than at 8 weeks postnatally, with 13.5% scoring above threshold for

probable depression at 32 weeks and 9.1% at 8 weeks postnatally (Evans et al., 2001). The study used self-report measures (Edinburgh Postnatal Depression Scale [EPDS] and Crown-Crisp Experiential Index [CCEI]) and did not confirm diagnoses of depression. The variation in rates found is probably a result of different populations studied. It should be noted that Gavin and colleagues (2005) used only studies where depression had been diagnosed according to recognised criteria rather than self-report measures. These authors concluded that it was not possible, given the currently available research, to state with any certainty whether there is a difference in rates between pregnancy trimesters or between months postnatally. But it was possible to say that all these studies are clear that pregnancy is not protective against depression.

Low mood after childbirth (sometimes called 'baby blues') is very common, occurring in 30 to 80% of women in the first weeks, but is usually mild and transient and needs to be differentiated from clinical depression in the postnatal period (Henshaw et al., 2003). There has been some debate over the putative increased incidence of depression in the postnatal period with early research reporting incidence to be raised approximately threefold in the first 5 weeks postnatally (Cox et al., 1993) and recent longitudinal population-based studies have observed increased incidence during the postnatal period (Ban et al., 2012; Munk-Olsen et al., 2006). Incident cases of depression in the postnatal period may reflect lack of identification or measurement of depression starting in pregnancy. Recent studies have found that at least a third of 'postnatal depression' begins in pregnancy or pre-pregnancy (Heron et al., 2004; Wisner et al., 2013).

As with depression at other times, depression in the postnatal period is often self-limiting within a few months, but around 30% of women remain unwell beyond the first year after childbirth and there is high risk (around 40%) of subsequent postnatal and non-postnatal relapse (Goodman, 2004; Cooper & Murray, 1995; Wisner et al., 2004a).

The *Confidential Enquiries into Maternal Deaths* (Oates & Cantwell, 2011) has consistently found a mental health problem to be one of the leading causes of maternal death in the UK, with over half of these deaths caused by suicide. In the last four enquiries over half of the women who died from suicide had a previous history of severe mental illness (affective psychosis or severe depressive illness); drug misuse is consistently reported in around a third of suicides (suicides during pregnancy remain relatively uncommon, and most occur following childbirth) (Oates & Cantwell, 2011). The majority of suicides in pregnant and postnatal women (about 60%) occur in the 6 weeks before, and the 12 weeks after, childbirth.

2.3.2 Anxiety disorders

The prevalence of most anxiety disorders in pregnancy and the postnatal period is similar to other times in women's lives; for example a large US population-based study found a 13% past-year prevalence of any anxiety disorder in currently pregnant or postnatal women, comparable to non-pregnant women (Vesga-Lopez et

al., 2008); the prevalence of anxiety symptoms is even higher (for example, Wenzel et al., 2003; Heron et al., 2004), particularly in pregnancy. For example, a large-scale community prospective study of around 8,300 women (based on the ALSPAC), which measured anxiety symptoms during pregnancy and the postnatal period (from 18 weeks' gestation to 8 months postnatally), found 14.6% scored above threshold at 18 weeks' gestation (a score of 9 or more on the anxiety items of the CCEI), while 8% scored above threshold at 8 weeks postnatally, with 2.4% *de novo* presentations (Heron et al., 2004). Two-thirds of women reporting anxiety during pregnancy also reported anxiety postnatally. Anxiety disorders are often comorbid with depressive disorders (NCCMH, 2011) and this seems to be particularly true for pregnant and postnatal women, with around two thirds of those with depression also having a comorbid anxiety disorder (Lydsdottir et al., 2014; Wisner et al., 2013).

A systematic review of anxiety disorders in pregnancy and the postnatal period (Ross & McLean, 2006) reported the prevalence of panic disorder at 1.3 to 2%, but there are few controlled studies to establish whether pregnancy is associated with reduced symptoms (which has been reported from some small studies) or whether panic disorder worsens in the postnatal period. A large US population-based study found a 13% past-year prevalence of any anxiety disorder in currently pregnant or postnatal women, comparable to non-pregnant women (Vesga-Lopez et al., 2008). There are even fewer data on GAD, but there is some emerging evidence suggesting higher rates in pregnancy with a reduction in the postnatal period, though these rates are still higher than those reported in general population studies (Buist et al., 2011; Ross & McClean, 2006). There is also a growing literature on a specific phobia, tokophobia (fear of childbirth), which may pre-date pregnancy (known as 'primary' tokophobia). Fear of childbirth may also be secondary to traumatic childbirth (sometimes referred to as 'secondary' tokophobia), but this may be more helpfully conceptualised as a trauma symptom or as part of a presentation of post-traumatic stress disorder (PTSD); symptoms may also be caused by another mental health problem, such as depression (Rouhe et al., 2011; Storksen et al., 2011). The prevalence of tokophobia is unclear – up to 80% of low risk pregnant women describe common childbirth anxieties, with 6 to 10% reporting pathological levels of fear (Saisto & Halmesmaki, 2003), but this includes women who do not fulfil diagnostic criteria for a specific primary phobia and therefore the prevalence is likely to be much lower. Fear of childbirth in pregnancy has been associated with an increased probability of having an emergency or elective Caesarean section in some studies (Ryding et al., 1998; Waldenström, 2006).

Other specific phobias of relevance to pregnancy include needle phobia, which can restrict pain relief options (such as an epidural during labour) for these women and lead to them refusing blood tests -- as a result medical conditions might go undetected, with potentially serious consequences (Oates & Cantwell, 2011).

Despite the view that anxiety disorders only constitute mild mental health problems, they are associated with significant disability and this, combined with the emerging

evidence of possible negative effects on the fetus, demonstrable in infancy, reinforces the view that more attention needs to be paid to these disorders.

A recent systematic review and meta-analysis of obsessive-compulsive disorder (OCD) reported overall prevalence estimates of 1.08% for women in the general population, 2.07% during pregnancy, and 2.43% during the postnatal period - pregnant or postnatal women are approximately 1.5 to 2 times more likely to experience OCD than the general population (Russell et al., 2013). The potential difference between pregnancy and the postnatal period should be viewed with caution because of the limited data available. However it appears reasonable to conclude that the risk of OCD is greater when women are pregnant or postnatal (Russell et al., 2013)—whether that risk is greater for postnatal compared with pregnant women requires further research.

Symptoms of PTSD following childbirth have been reported in a number of women. A review of links between childbirth and PTSD in women following a live birth found prevalence figures for a ' PTSD-profile' (that is, symptom criteria of DSM-IV B, C and D) of between 2.8 and 5.6% at around 6 weeks postnatally, which reduced to 1.5% by 6 months postnatally (Olde et al., 2006). A more recent systematic review and meta-analysis reported rates of 3.1% across community samples (that is, nontargeted postnatal women) in studies considering rates 1- 18 months postnatally where symptoms relate specifically to childbirth (Grekin & O'Hara, 2014). This is consistent with the usual course of PTSD, which appears to have a high remittance rate following the index traumatic event (NCCMH, 2005). The rate in studies using DSM-IV criteria was between 1.7% (1 to 13 months postnatally) and 2.8% (6 months postnatally). Czarnocka and Slade (2000), in a self-report questionnaire study, found that 3% of their sample of 264 women showed clinically significant levels on all three PTSD dimensions and 24% on at least one dimension. However, most studies underestimate the total prevalence of PTSD in the postnatal period by examining PTSD related to traumatic childbirth experiences only; higher rates are observed in pregnancy when diverse trauma experiences are included (point prevalence 6.8%) (Seng et al., 2010). PTSD in pregnancy and the postnatal period is also highly comorbid with depression (Seng et al., 2010). Stillbirth has also been identified as a stressor for PTSD symptoms during a subsequent pregnancy (Turton et al., 2001), as has premature delivery.

2.3.3 Eating disorders

Anorexia nervosa in pregnant women is less common than in the general population because of the reduced fertility and fecundity associated with this disorder and its usual onset in adolescence. In a follow-up study of people with anorexia nervosa (n=140), fertility was reduced to one third of the expected rate (Brinch et al., 1988). However, pregnancy does occur in women with anorexia nervosa; pregnancy in women with bulimia nervosa is less rare since this disorder is less likely to cause infertility, although as many as 50% may experience amenorrhoea or oligo-amenorrhoea (Fahy & Morrison, 1993) at some point in the course of the illness. Oligoamenorrhoea or vomiting oral contraceptives may increase the risk of

unplanned pregnancy among women with bulimia nervosa (Morgan et al., 1999). Recent research suggests that around 5 to 7.5% of pregnant women may meet diagnostic criteria for an eating disorder (Easter et al., 2013; Watson et al., 2013). There is also preliminary evidence that pregnancy can lead to remission from bulimia nervosa but worsen symptoms of binge eating disorder (Watson et al., 2013).

There is little research into eating disorders in the postnatal period but onset or recurrence of eating disorders can occur (Stein et al., 1996), disordered eating persists in a substantial proportion of women meeting criteria for either full or subthreshold eating disorders (Knoph et al., 2013) and is associated with weaning difficulties. Eating disorders are also associated with an increased risk of depression and anxiety in pregnancy and the postnatal period (Micali et al., 2011).

2.3.4 Psychotic disorders (schizophrenia and bipolar disorder)

Psychosis is defined as a mental disorder which is characterised by hallucinations and/or delusions and related symptoms with severe functional impairment. Bipolar disorder is characterised by depression and mania (abnormally elevated mood or irritability and related symptoms with severe functional impairment). Although women with psychotic disorders are less fertile than the general population (Howard et al., 2002), recent changes in the types of antipsychotic medications prescribed (with consequent reductions in the prevalence of hyperprolactinaemia, which impacts on fertility) has led to less severe subfertility (Vigod et al., 2012), particularly for women with bipolar disorder, with adolescents having higher fertility than the general population (Vigod et al., 2014). Pregnant women with psychotic disorders are particularly likely to have risk factors for physical health problems (see Section 2.3.8).

There are limited data on the prevalence and incidence of psychotic disorders in pregnancy, but although prevalence appears to be similar to that found in non-pregnant women of childbearing age, the incidence of first psychiatric admissions is lower (Munk-Olsen et al., 2006). It has recently been recognised that symptoms of depression in pregnancy and the postnatal period may actually constitute an underlying bipolar disorder; recent studies have found rates of 13% for bipolar II disorder (bipolar disorder without psychosis) in women with high levels of depressive symptoms in pregnancy (Lydsdottir et al., 2014) and rates of 22% in the postnatal period (Wisner et al., 2013).

Most women with a psychotic disorder have children at some point in their lives (Howard et al., 2001) and there is mixed evidence on the risk of relapse in pregnancy for these women. Prospective cohort studies suggest there is an increased risk of relapse in pregnant women with bipolar disorder who discontinue prophylactic medication such as mood stabilisers (Viguera et al., 2007), but there is little evidence on the course of schizophrenia in pregnancy. In the postnatal period, psychosis is associated with an increased risk of relapse - this is particularly notable for bipolar disorder and both retrospective and population registry studies suggest that women with bipolar disorder have at least a 1 in 5 risk of having a severe recurrence

following childbirth (Di Florio et al., 2013; Jones et al., 2005; Munk-Olsen et al., 2009) and a higher risk (around 1 in 2) of experiencing any mood episode in the postnatal period including depression (see below). This increased risk of relapse occurs in the first few months after childbirth for women with bipolar disorder; by contrast women with schizophrenia are at an increased risk, but of lower magnitude, throughout the first postnatal year (Munk-Olsen et al., 2006; Munk-Olsen et al., 2009).

2.3.5 Postpartum psychosis

Psychosis in the early postnatal period (up to 3 months after childbirth) is often termed postpartum or puerperal psychosis (this guideline uses the term 'postpartum psychosis'). Whether it is a distinct diagnosis has been the subject of considerable debate, but most commonly it takes the form of mania, severe depression, or a mixed episode with features of both high and low mood. DSM-V does not categorise postpartum psychosis as a separate entity and uses a perinatal-onset specifier (that is, pregnancy or up to 4 weeks after childbirth), while ICD-10 has a special category (though advises against its use). However, research has consistently reported an increase in rates of psychosis in the first 90 days after childbirth, with 21-fold higher rates of inpatient admission in this period compared with other times, with figures of around 1 per 1000 (Kendell et al., 1987; Munk-Olsen et al., 2006).

The incidence of postpartum psychosis is also unclear, partly because many studies include episodes of bipolar disorder that may not have been psychotic (Harlow et al., 2007). The incidence rate commonly quoted is 1 to 2 per 1000 deliveries, although it has been suggested that if more stringent criteria are applied, such as admission with definite psychotic symptoms within 2 weeks of childbirth, the rate is between 0.5 and 1 per 1000 deliveries (Kumar, 1989; Terp & Mortensen, 1998). A later study of 502,767 first-time mothers found an average rate of 0.68 per 1000 (Nager et al., 2005). This study excluded those with an admission for psychotic disorder within 2 years before childbirth. This would have removed those with existing severe mental illness, such as bipolar disorder, liable to relapse and thus indicates that childbirth is a risk factor for the onset of psychosis, albeit a very small one.

Postpartum psychosis is characterised by sudden onset and rapid deterioration and the clinical picture often changes rapidly, with wide fluctuations in the intensity of symptoms (which commonly include delusions and hallucinations, and confusion or perplexity) and severe mood swings. Most episodes of postpartum psychosis start within 2 weeks of childbirth, with retrospective accounts suggesting that symptoms began in the first few postnatal days or even during labour (Heron et al., 2008) but the increased risk appears to persist to some extent for the first 3 months after childbirth (Valdimarsdóttir et al., 2009). Women with a history of a previous postpartum psychosis are at very high risk with greater than 1 in 2 deliveries affected (Robertson et al., 2005) and for women with bipolar disorder, a family history of bipolar disorder or postpartum psychosis gives a similarly high risk in the postpartum period (Jones et al., 2001; Munk-Olsen et al., 2007). However, many

(around 50%) women have no history that indicates they are at high risk (Valdimarsdóttir et al., 2009)

2.3.6 Drug and alcohol-use disorders

Drug and alcohol misuse in pregnancy are markers of complex pregnancies, multiple comorbidities and adverse obstetric fetal and infant outcomes, and are often associated with limited access to healthcare during pregnancy. In 2006-8, women who misused drugs accounted for 11% of all maternal deaths and 31% of maternal deaths from suicide; 44% received little or no healthcare during pregnancy (Oates & Cantwell, 2011). Women who misuse alcohol and drugs are more likely to smoke than other pregnant women (smoking is the leading preventable cause of fetal and infant adverse outcomes in the UK [Royal College of Physicians, 2010]) and have significant other complex problems including poor diet, poverty and domestic violence and abuse, which are also associated with adverse maternal and child outcomes. Postnatally, alcohol and drug misuse are significantly associated with sudden infant death syndrome and an adverse impact on parenting. Many women stop using alcohol or other drugs once they know they are pregnant but relapse is common.

Alcohol misuse

In 2010, two in five mothers (40%) reported drinking some alcohol during pregnancy (fewer than the 54% in 2005). Mothers aged 35 or over (52%), mothers from managerial and professional occupations (51%) and mothers from a white ethnic background (46%) were more likely to report drinking during pregnancy (McAndrew et al., 2010). Among women who drank during pregnancy, consumption levels were low. Only 3% of all expectant mothers drank more than two units of alcohol per week on average; however these data are likely to be an underestimate of drinking behaviour as women are aware that current advice is to avoid alcohol. Around 10% of women childbearing age are binge drinkers and are likely to have consumed potentially harmful levels of alcohol before they knew they were pregnant. Binge drinking before pregnancy is a strong predictor of both drinking during pregnancy and binge drinking during pregnancy (Ethen et al., 2009).

Alcohol is teratogenic and there is some debate on the safe limit of alcohol use in pregnancy due to the difficulty in establishing effects of low to moderate levels of drinking in observational studies (Henderson et al., 2007; Gray et al., 2009). There is therefore insufficient evidence to define any threshold for low-level drinking in pregnancy. However there is well established evidence that high levels of alcohol consumption are associated with infertility, miscarriage, preterm labour, stillbirths and a spectrum of behavioural and neurocognitive impairments (known as 'alcohol related neurodevelopmental disorder') in the developing fetus (O'Leary et al., 2009); the most severe end of the spectrum is 'fetal alcohol syndrome' (a triad of dysmorphic facial features, impaired growth and central nervous system abnormalities), which occurs in around 0.21 per 1000 live deliveries in the UK (Department of Health, 2002).

Illicit drug misuse

There are no national estimates for pregnant women who misuse drugs in the UK, but studies report that approximately a third of drug users in treatment are female and over 90% of these women are of childbearing age (15–39 years of age). It has been estimated that 200,000 to 300,000 children in England and Wales have one or both parents with a serious drug problem[1]. Inner city maternity services report around 10 to 15% of pregnant women with positive drug screens, mostly cannabis (Sherwood et al., 1999; Williamson et al., 2006), and polydrug misuse is common (Mayet et al., 2008). Drugs readily cross the placenta and are associated with adverse pregnancy outcomes including stillbirth, prematurity, and low birthweight babies (Mayet et al., 2008). Opioids are particularly associated with neonatal withdrawal syndrome (Patrick et al., 2012) and neurobehavioural problems, increased neonatal mortality and sudden infant death syndrome (Amato et al., 2013).

2.3.7 Personality disorder

There has been little research into personality disorder in pregnancy and the postnatal period. In a recent survey in England, around 1.4% of women aged 16 to 35 years had a diagnosis of borderline personality disorder and 0.4% had antisocial personality disorder (McManus et al., 2009). Although there are no studies in maternity populations in the UK, a Swedish study reported that 6% of women of childbearing age had a personality disorder (Borjesson et al., 2005), although this study used a self-report measure and did not report the prevalence of individual personality disorders. Severe personality disorder is associated with disturbances in mother-infant interaction (for example, Hobson et al., 2009) and loss of custody (Howard et al., 2003).

2.3.8 Physical health problems

Women with a mental health problem in pregnancy and the postnatal period have a higher prevalence of risk factors for physical health problems compared with pregnant and postnatal women without a mental health problem. These include smoking, nutritional deficits, obesity, hypertension and domestic violence (Royal College of Physicians & Royal College of Psychiatrists 2013; McColl et al., 2013; Molyneaux et al., 2014; Katon et al., 2012; Boden et al., 2012a) which can lead to physical health problems for the mother and adverse outcomes for the fetus. In addition, symptoms of medical conditions such as eclampsia, infection or pulmonary embolus may be misattributed to a mental health problem and this has led to deaths in new mothers (Oates & Cantwell, 2011).

[1] Working group report from the Advisory Council on the Misuse of Drugs (2003). Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/120620/hidden-harm-full.pdf

2.4 AETIOLOGY OF MENTAL HEALTH PROBLEMS IN PREGNANCY AND THE POSTNATAL PERIOD

The variation in the presentation, course and outcomes of mental health problems in pregnancy and the postnatal period is reflected in the breadth of theoretical explanations for their aetiology, including genetic, biochemical and endocrine, psychological and social factors. As already discussed most mental health problems are not unique to pregnancy and the postnatal period and the aetiological factors involved will reflect the aetiology of mental health problems at other times in women's lives, which include a history of psychopathology, psychosocial adversity, childhood and adulthood abuse, and social support (Lancaster et al., 2010; Howard et al., 2013; Robertson et al., 2004; Ross & Dennis, 2009). As for specific factors connected to pregnancy and the postnatal period, the predominant specific hypothesis has been that hormonal changes (including thyroid and pituitary hormones, cortisol and gonadal hormones) might be important, but no clear aetiological association has emerged (Hendrick et al., 1998). Nevertheless there is evidence of familiality of the trigger for postpartum psychosis (Jones et al.2001) and of a 'reproductive subtype' of depression characterised by a particular sensitivity to changes in reproductive hormones (Bloch et al., 2000), increased risk of premenstrual, postnatal and perimenopausal depression (Buttner et al., 2013; Murray et al., 1996), and a personal or family history of depression in the postnatal period (Craig, 2013). Specific traumas including stillbirth, infant complications and other forms of traumatic childbirth experiences are associated with mental health problems, particularly PTSD (Adeyemi et al., 2008; Andersen et al., 2012; Furuta et al., 2012; Turton et al., 2001). Maternity populations increasingly have significant proportions of women who were not born in the UK and there is emerging evidence that refugees, asylum seekers and trafficked pregnant women are at increased risk of mental health problems (Collins et al., 2011; Oram et al., 2012).

2.5 TREATMENT IN THE NHS

In common with mental health problems at other stages in people's lives, detection in pregnancy and the postnatal period by different professionals is variable, and this inevitably results in under-treatment. Stigma and concerns about potential statutory involvement in the care of the baby may add to the reluctance to seek help, even where it is recognised by the woman herself. The detection of mental health problems in pregnancy and the postnatal period is the subject of Chapter 5 and will not be discussed in detail here. However, the identification of depression and anxiety in the general population gives an indication of the consequences of under detection. For example of the 130 depressed people per 1000 population, only 80 will consult their GP. Of these 80 people, 49 are not recognised as depressed, mainly because most such patients are consulting for a somatic symptom and do not consider themselves mentally unwell, despite the presence of symptoms of depression (Kisely et al., 1995). This group also has milder illnesses (Goldberg et al., 1998; Thompson et al., 2001). GPs and other non-mental health specialists vary in their ability to recognise depressive illnesses, with some recognising the vast majority of the patients found to be depressed at independent research interview

and others recognising very few (Goldberg & Huxley, 1992; Üstün & Sartorius, 1995).

The communication skills of healthcare professionals make a vital contribution to determining their ability to detect emotional distress, and those with superior skills allow their patients to show more evidence of distress during their interviews, thus making detection easy. Those with poor communication skills are more likely to collude with their patients, who may not themselves wish to complain of their distress unless they are asked directly about it (Goldberg & Bridges, 1988; Goldberg et al., 1993).

In summary, those with severe mental illness, and those presenting with psychological symptoms, are especially likely to be recognised, while those presenting with somatic symptoms for which no cause can be found are less likely to be recognised. It is probable that the position described above for depression holds for most, if not all, mental health problems. In pregnancy and the postnatal period, women are in frequent contact with healthcare professionals, which provides opportunities for increasing healthcare professionals' awareness of mental health problems and improving their detection skills.

2.5.1 The provision of care for mental health problems in pregnancy and the postnatal period in the NHS in England and Wales

The large majority of women (over 90%) with mental health problems in pregnancy and the postnatal period are treated in primary care, where most common mental health problems (depression and anxiety disorders) are treated. The remainder receive care from specialist mental health services, including general adult services, liaison services and specialist perinatal services. Provision of specialist perinatal mental health services is covered in Chapter 4.

2.5.2 Psychological interventions

There is little evidence, other than in the treatment of depression, on the differential effectiveness of psychological interventions during pregnancy and the postnatal period. The major difference is the shifting risk-benefit ratio, relating to the possible risks associated with the use of psychotropic medication (see below). For example, in the NICE depression guideline (NICE, 2009a) antidepressants are recommended for the treatment of moderate depression, but in pregnancy and the postnatal period the threshold for the use of psychotropic medication will be higher, and access to psychological interventions may need to be expedited. Given the limited availability of psychological treatments, even with the advent of the Improving Access to Psychological Therapies (IAPT) programme, this may present a considerable challenge for perinatal services.

2.5.3 Pharmacological interventions

As with psychological interventions, there is little evidence to suggest that pharmacological treatments (the mainstay of treatment of mental health problems in

the NHS) have any differential benefit in pregnancy or the postnatal period from their use in other adult populations. As stated above, the major difference is in the shifting risk-benefit ratio in pregnancy and the postnatal period. This relates to the possibility of increased teratogenic and neurodevelopmental risks to fetus (associated with the use of psychotropic medication. The potential risks, which are not clear (see chapter 8) need to be balanced carefully in the case of each woman and set against the baseline risks of malformation, the likely benefits of any treatment and the risks of untreated mental health problems that increase the baseline risk of malformations. Clinicians also need to be aware of potential changes in the pharmacokinetics of drugs in pregnant women due to increased fluid balance, particularly in the third trimester. Women may also be less able to tolerate some side effects during pregnancy or the postnatal period.

2.5.4 The organisation of perinatal mental health services

The organisation of perinatal services does not follow any consistent pattern across England and Wales; provision is variable, recommendations from various sources are often not coordinated (Department of Health, 2004; Department of Health, 2002; Mann, 1999), but there are now commissioning guidelines for perinatal mental health services. The service structures required to support effective mental healthcare in pregnancy and the postnatal period are discussed in Chapter 4.

One challenge faced by those involved in the care of women with mental health problems in pregnancy and the postnatal period is the wide range of services that women use at this time. This requires close communication and agreed plans of care at the level of the individual woman and for effective collaborative working arrangements at a service level between primary care (GP, health visitor, psychological therapy services [IAPT programme] and counsellor), maternity services (midwife and obstetrician) and, where appropriate, secondary care mental health services and also social services and the independent and voluntary sectors. This network of care must not only consider the needs of the woman and her child but also other family member and carers. Poor communication has often been identified as the reason for poor-quality care and was behind the development of the care programme approach in the UK healthcare system (Department of Health, 1999b; Department of Health, 2008).

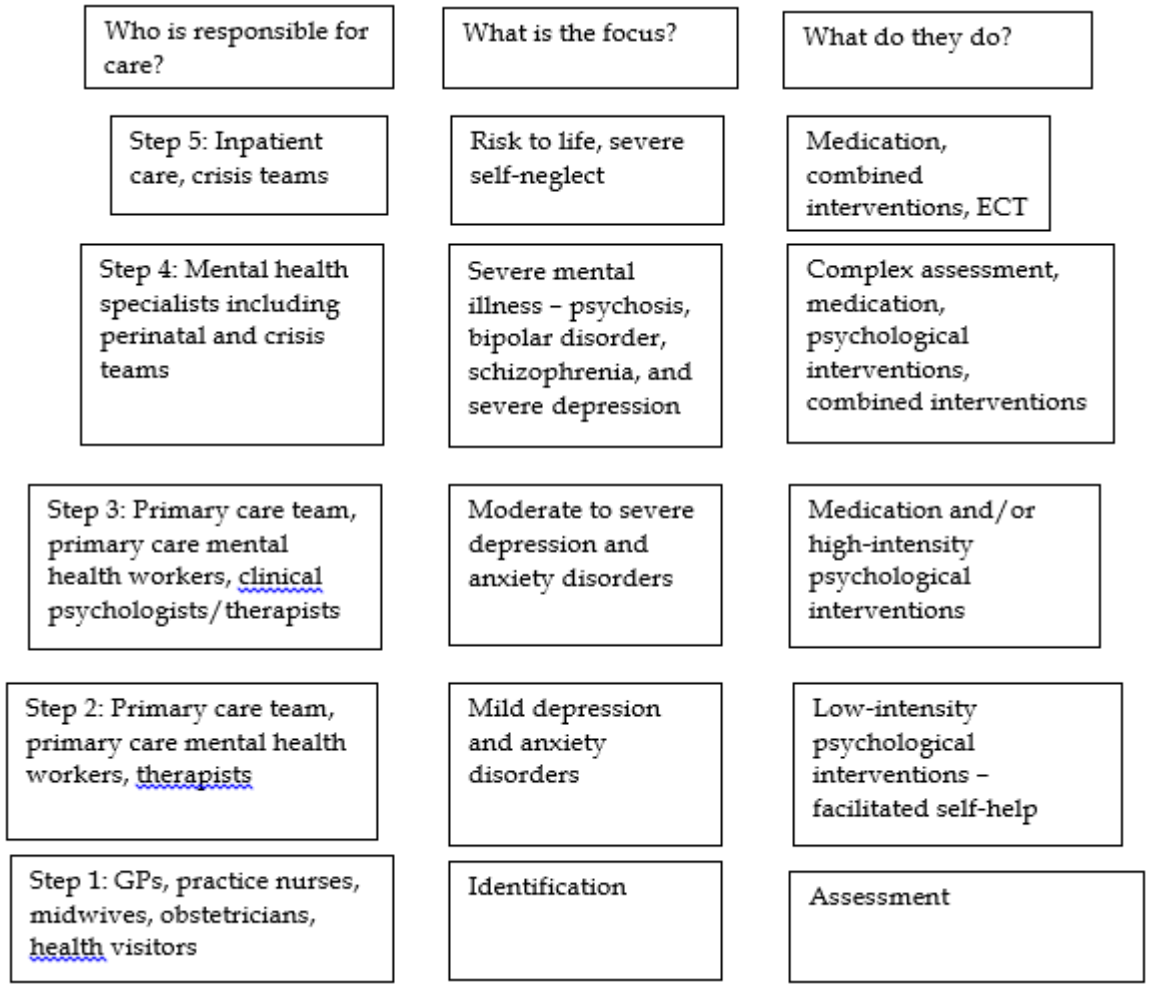
In addition to providing effective communication, services need to be organised in ways that promote the development of cost-effective treatments and provide clear pathways, which are understandable to both providers and recipients of care. The experience for the individual woman of the involvement of multiple professionals can be bewildering and overwhelming. If not properly coordinated to prevent duplication, overlaps and gaps in service, this may also be counter-therapeutic. Despite the involvement of multiple services, it can be women's experience that their needs for practical help at this critical time are neglected because services tend to emphasise processes of assessment, monitoring, psychotherapeutic intervention and

medication but rarely address the practical demands of looking after one or more young children day and night while mentally unwell.

In a number in the NICE guidelines, a 'stepped' or 'tiered care' model of service delivery has been developed, which draws attention to the different needs that women with mental health problems in pregnancy and the postnatal period have, depending on the characteristics of their problem and their personal and social circumstances, and the responses that are required from services. This stepped/tiered model is a hybrid of two ideas. At one end, is 'pure' stepped care where people are offered the least intrusive and lowest intensity intervention likely to be effective in helping them. They would only receive a more intensive, or complex, intervention if their symptoms did not improve at an earlier step. At the other end, there is stratified care where often the intervention is linked to a particular diagnosis or service provider. Patients are directed to the service or professional who is seen to provide the optimum intervention for that person. Both these models are sometimes 'overlaid' onto a service model that identifies various tiers of services often provided by different organisations. The model also assumes effective working relationships across the system; for example, a specialist mental health or perinatal service may provide advice, training or consultation on the management of patients at levels one and two.

There are advantages and disadvantages to each of these models. The following is a model that attempts to outline the relationship between severity of illness and the most appropriate professional skill set in the corresponding organisational structure (see Figure 1).

Figure 1: The stepped/tiered care model



2.6 THE ECONOMIC COSTS OF MENTAL HEALTH PROBLEMS IN PREGNANCY AND THE POSTNATAL PERIOD

Existing evidence on the financial implications of the presence of mental health problems in pregnancy or in the first postnatal year is very limited. A systematic review of the literature identified two UK-based studies. One study was conducted in 2002 and looked at the health and social care costs of depression in the postnatal period; and another more recent study looked at the costs associated with paternal depression. The review also identified three international studies (that is, from US, Canada and Australia) that explored the additional healthcare resource use and/or financial costs associated with care of women with depression in the postnatal period and their babies. No studies examining the economic burden imposed by women with other mental health problems in pregnancy and the postnatal period were found in the literature. The existing evidence on financial costs associated with substance misuse in pregnancy is only from North America.

Petrou and colleagues (2002) estimated the health and social service costs of depression in the postnatal period in a cohort of 206 women at high risk of developing the condition. The study was conducted in Reading, UK, between 1997 and 1999. Women were identified as being at high risk using a predictive index for depression in the postnatal period. Costs were estimated for participating women and their babies over 18 months after childbirth and included costs of inpatient, outpatient, day care and community services. Paediatric and childcare services were recorded separately. The mean mother–infant costs over 18 months were found to be £3,647 when women developed depression in the postnatal period (as measured by the Structured Clinical Interview for DSM Disorders – II [SCID-II]) and £3,056 when no depression was diagnosed (uplifted to 2013 prices). The overall cost difference between the two groups was £591 (p=0.17). Also, the community care costs for women with depression in the postnatal period were higher compared with respective costs for women without depression in the postnatal period (p=0.01). The authors estimated that, with approximately 700,000 women giving birth in the UK annually and a 13% incidence of depression in the postnatal period, the economic burden of this condition to the health and social services in the UK amounted to roughly £54 million annually (range £52 to £65 million). It was acknowledged that this value might in reality be a conservative estimate, given that the condition was likely to have longer-term consequences in terms of health status and health service utilisation over the woman's and her child's lifetime and in terms of the child's educational requirements. Moreover, with evidence that women not at high risk for depression in the postnatal period had fewer contacts in pregnancy and the postnatal period than the study population, the additional costs associated with care of women developing depression in the postnatal period might be even higher in comparison to respective costs associated with care of the population of women giving birth as a whole.

Similarly, in the recent report prepared for the Post and Antenatal Depression Association (PANDA) in Australia (PANDA, 2012) the financial costs associated with maternal depression in pregnancy and the postnatal period were estimated. The study included direct healthcare costs relating to primary care, psychiatrist and allied health services, medications, hospitals and community services. Total direct healthcare costs of maternal depression in the postnatal period for the annual cohort of 70,997 were estimated to be AU$61 million; no data were available for depression during pregnancy. The highest cost category was hospital services, which were estimated to be AU$40 million. The next most significant categories were psychiatrist and allied health services (AU$8 million), primary care (AU$6 million), community mental health services (AU$4 million) and medications (AU$4 million). The authors also estimated the cost of lost productivity to be AU$87 for maternal depression during pregnancy. The additional costs associated with government expenditure on health and related services that were provided to people with depression in pregnancy were estimated to be AU$45 million.

In Minnesota in the US, Dagher and colleagues (2013) examined the association between depression in the postnatal period and healthcare expenditure 11 weeks

after childbirth in a sample of employed women (n=638) from three community hospitals in 2001. The mean costs from childbirth until 11 weeks postnatally were found to be US$1,046 in women who developed depression in the postnatal period and US$365 when no depression was diagnosed (2001 prices). The overall cost difference between the two groups was US$681 ($p < 0.001$). In another study, O'Brien and colleagues (2009) estimated the costs of untreated depression in pregnancy in Ontario, Canada. The authors estimated that in 2006–07 approximately 2,593 women who discontinued their antidepressants had a depressive relapse. This resulted in maternal healthcare costs of approximately CA$1 million and the cost of caring for preterm babies of women with depression in the first year after childbirth was estimated to be CA$9 to CA$13 million. Also, there is evidence that women with depression in the postnatal period are less likely to attend scheduled appointments and are more likely to present to more expensive accident and emergency departments (Minkowitz et al., 2005; Stock et al., 2013).

The mental health needs of fathers/partners whose health and functioning will inevitably be affected by mental health problems in women, are also important and should be considered. In the UK, Edoka and colleagues (2011) estimated healthcare costs of paternal depression in the postnatal period using self-reported resource-use data collected alongside longitudinal study. The authors collected data on healthcare resource use over the first postnatal year from 192 fathers recruited from two postnatal wards in southern England. Three groups of fathers were identified: fathers with depression (n=31), fathers at high risk of developing depression (n=67) and fathers without depression (n=94). The mean father–infant costs were estimated at £1,104, £1,075 and £945 (2008 prices) in these three groups, respectively ($p=0.796$). Moreover, after controlling for potentially confounding factors, paternal depression was associated with higher community care costs (mean cost difference of £132; $p=0.005$). Within this category, increased contacts with GPs and psychologists made the highest contribution to the observed cost difference between those with and without depression.

No studies examining the economic burden imposed by women with other mental health problems in pregnancy and the postnatal period were found in the literature. However, some studies report that women with eating disorders are more likely to have delivery by Caesarean section. Similarly fear of childbirth in pregnancy has been associated with an increased risk of costly emergency Caesarean sections.

There is a bit more evidence on financial costs associated with substance misuse in pregnancy; however, it is mainly from North America. In Canada, Popova and colleagues (2014) estimated the number of children (0-18 years) in care with fetal alcohol syndrome spectrum disorders and looked at the associated costs by age group, gender, and province/territory in 2011. The estimated number of children in care with fetal alcohol syndrome spectrum disorders ranged from 2,225 to 7,620, with an annual cost of care ranging from CA$58 to CA$198 million. The highest overall cost (CA$30 to CA$101 million) was for 11-15 year-olds. Similarly, in another study Popova and colleagues (2013) estimated the utilisation of specialised addiction

treatment services and the associated cost for people with fetal alcohol syndrome spectrum disorders. This was a modelling study with data obtained from various national sources. The cost of specialised addiction treatment services for people with fetal alcohol syndrome spectrum disorders in Canada in 2010-11 ranged from CA$2 to CA$4 million, based on 5,526 outpatient visits and 9,529 resident days. When the sensitivity analysis was performed, the cost of specialised addiction treatment services ranged from approximately CA$1 to CA$5 million. In another Canadian study, Stade and colleagues (2009a) estimated the annual cost associated with fetal alcohol syndrome spectrum disorders at the individual level to be CA$21,642 (95% confidence interval [CI], CA$19,842 to CA$24,041) and the cost of fetal alcohol spectrum disorders annually to Canada from day of birth to 53 years old, was estimated to be CA$5 billion (95% CI, CA$4.12 to CA$6.4 billion). These data do not include the cost of children in care of child protection systems, special education, costs to the justice system or supportive housing or addictions treatment. Brownell and colleagues (2013) examined health, education and social service use of individuals with fetal alcohol spectrum disorders in Canada. The authors used a matched-cohort design of health, education and social service data that were linked with clinical records on individuals 6+ years diagnosed with fetal alcohol spectrum disorders between 1999-2000 and 2009-2010. Matching was done with a general population and asthma group by age, sex and area-level income. Hospitalisations were higher in the fetal alcohol spectrum disorders group compared with the general population and asthma group, and physician visits and overall prescriptions in the fetal alcohol spectrum disorders group differed from only the general population group. Antibiotics, pain killers and antipsychotics were similar across all groups whereas antidepressants and psychostimulants were higher in the fetal alcohol spectrum disorders group. Also, attention deficit hyperactivity disorder (ADHD) was higher in the fetal alcohol spectrum disorders group. Education and social service use was higher for the fetal alcohol spectrum disorders group than either of the other groups for all measures (that is, grade repetition, receipt of any special education funding, family receipt of income assistance, child in care, and receipt of child welfare services). In the US, Amendah and colleagues (2011) examined medical expenditures of children with fetal alcohol spectrum disorders. Children with fetal alcohol spectrum disorders incurred annual mean medical expenditures that were nine times as high as those of children without disorder during 2005 (US$16,782 versus US$1,859). In another US study, Kalotra and colleagues (2002) reviewed literature pertaining to the costs related to the birth of a drug and/or alcohol exposed baby. Total lifetime costs for caring for those children that survive ranged from US$750,000 to US$1 million.

As regards neonatal abstinence syndrome, Patrick and colleagues (2012) conducted a retrospective analysis of a nationally representative sample of newborn babies with neonatal abstinence syndrome between 2000 and 2009. In 2009, newborn babies with neonatal abstinence syndrome were more likely than all other hospital births to have low birthweight and respiratory complications. Mean hospital charges for discharges with neonatal abstinence syndrome was US$53,400 (95% CI, US$49,000 to US$57,700) in 2009 (in 2009 prices). Similarly, Backes and colleagues (2012) conducted a

retrospective review (2007-09) of babies born to mothers maintained on methadone in an antenatal drug misuse programme. The average hospital cost for each baby ranged from US$13,817 to US$27,546. Smith and colleagues (2002) report that substance misuse compromises appropriate parenting practices and increases the risk of child maltreatment. Costs of service provision for looked after children impose great economic burden on healthcare and social care services in England. It has been estimated that in the 2009-10 financial year around £3 billion were spent on looked after children's services in England. This equates to £37,669 per looked after child per annum in 2009-10 (Harker, 2012).

Besides the costs reported in the above studies, other factors associated with the care of babies born to mothers with mental health problems or those with drug or alcohol-use disorders in pregnancy need to be considered. There is evidence of increased risk of adverse outcomes for these mothers' children including depression, conduct disorder and anxiety disorders. The costs to society of these disorders are very high (Scott et al., 2001; King et al., 2006). Similarly, substance misuse during pregnancy can cause a range of physical and intellectual disabilities in the children of these mothers. These disabilities, in most cases multiple, can be extremely challenging to manage, they affect an individual for the rest of their lives and impose a substantial burden on health and social care services, and society as a whole.

3 METHODS USED TO DEVELOP THIS GUIDELINE

3.1 OVERVIEW

The development of this guideline followed *The Guidelines Manual* (NICE, 2012a). A team of health and social care professionals, lay representatives and technical experts known as the Guideline Development Group (GDG), with support from the NCCMH staff, undertook the development of a person-centred, evidence-based guideline. There are seven basic steps in the process of developing a guideline:

1. Define the scope, which lays out exactly what will be included (and excluded) in the guidance.
2. Define review questions that cover all areas specified in the scope.
3. Develop a review protocol for each systematic review, specifying the search strategy and method of evidence synthesis for each review question.
4. Synthesise data retrieved, guided by the review protocols.
5. Produce evidence profiles and summaries using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) system.
6. Consider the implications of the research findings for clinical practice and reach consensus decisions on areas where evidence is not found.
7. Answer review questions with evidence-based recommendations for clinical practice.

The clinical practice recommendations made by the GDG are therefore derived from the most up-to-date and robust evidence for the clinical and cost effectiveness of the interventions and services covered in the scope. Where evidence was not found or was inconclusive, the GDG discussed and attempted to reach consensus on what should be recommended, factoring in any relevant issues. In addition, to ensure a service user and carer focus, the concerns of service users and carers regarding health and social care have been highlighted and addressed by recommendations agreed by the whole GDG.

3.2 THE SCOPE

Topics are referred by the Secretary of State and the letter of referral defines the remit, which defines the main areas to be covered (see *The Guidelines Manual* [NICE, 2012a] for further information). The NCCMH developed a scope for the guideline based on the remit (see Appendix 1). The purpose of the scope is to:

- provide an overview of what the guideline will include and exclude
- identify the key aspects of care that must be included

- set the boundaries of the development work and provide a clear framework to enable work to stay within the priorities agreed by NICE and the National Collaborating Centre, and the remit from the Department of Health/Welsh Assembly Government
- inform the development of the review questions and search strategy
- inform professionals and the public about expected content of the guideline
- keep the guideline to a reasonable size to ensure that its development can be carried out within the allocated period.

An initial draft of the scope was sent to registered stakeholders who had agreed to attend a scoping workshop. The workshop was used to:

- obtain feedback on the selected key clinical issues
- identify which population subgroups should be specified (if any)
- seek views on the composition of the GDG
- encourage applications for GDG membership.

The draft scope was subject to consultation with registered stakeholders over a 6-week period. During the consultation period, the scope was posted on the NICE website (www.nice.org.uk). Comments were invited from stakeholder organisations. The NCCMH and NICE reviewed the scope in light of comments received, and the revised scope was signed off by NICE.

3.3 THE GUIDELINE DEVELOPMENT GROUP

During the consultation phase, members of the GDG were appointed by an open recruitment process. GDG membership consisted of: professionals in psychiatry, clinical psychology, nursing, health visiting, obstetrics, midwifery and general practice; academic experts in psychiatry and psychology, a mother infant specialist service users and a representative from a service user organisation. The guideline development process was supported by staff from the NCCMH, who undertook the clinical and health economic literature searches, reviewed and presented the evidence to the GDG, managed the process, and contributed to drafting the guideline.

3.3.1 Guideline Development Group meetings

Twelve GDG meetings were held between 14 March, 2013, and 2 September, 2014. During each day-long GDG meeting, in a plenary session, review questions and clinical and economic evidence were reviewed and assessed, and recommendations formulated. At each meeting, all GDG members declared any potential conflicts of interest (see Appendix 2), and service user concerns were routinely discussed as a standing agenda item.

3.3.2 Topic groups

The GDG divided its workload along clinically relevant lines to simplify the guideline development process, and GDG members formed smaller topic groups to undertake guideline work in that area of clinical practice. Topic Group 1 covered questions relating to case identification. Topic Group 2 covered psychological and psychosocial interventions, and Topic Group 3 covered pharmacological interventions. These groups were designed to efficiently manage the large volume of evidence appraisal prior to presenting it to the GDG as a whole. Each topic group was chaired by a GDG member with expert knowledge of the topic area (one of the healthcare professionals). Topic groups refined the review questions and the clinical definitions of treatment interventions, reviewed and prepared the evidence with the systematic reviewer before presenting it to the GDG as a whole, and helped the GDG to identify further expertise in the topic. Topic group leaders reported the status of the group's work as part of the standing agenda. They also introduced and led the GDG's discussion of the evidence review for that topic and assisted the GDG Chair in drafting the section of the guideline relevant to the work of each topic group.

3.3.3 Service users

Individuals with direct experience of services gave an integral service-user focus to the GDG and the guideline. The GDG included a service user and representatives of a national service user group. They contributed as full GDG members to writing the review questions, providing advice on outcomes most relevant to service users, helping to ensure that the evidence addressed their views and preferences, highlighting sensitive issues and terminology relevant to the guideline, and bringing service user research to the attention of the GDG. In drafting the guideline, they reviewed the chapter on experience of care and identified recommendations from the service user perspective.

3.3.4 Special advisors

Special advisors, who had specific expertise in one or more aspects of treatment and management relevant to the guideline, assisted the GDG, commenting on specific aspects of the developing guideline and making presentations to the GDG. Appendix 3 lists those who agreed to act as special advisors.

3.3.5 National and international experts

National and international experts in the area under review were identified through the literature search and through the experience of the GDG members. These experts were contacted to identify unpublished or soon-to-be published studies, to ensure that up-to-date evidence was included in the development of the guideline. They informed the GDG about completed trials at the pre-publication stage, systematic reviews in the process of being published, studies relating to the cost effectiveness of treatment and trial data if the GDG could be provided with full access to the complete trial report. Appendix 5 lists researchers who were contacted.

3.4 REVIEW PROTOCOLS

Review questions drafted during the scoping phase were discussed by the GDG at the first few meetings and amended as necessary. The review questions were used as the starting point for developing review protocols for each systematic review (described in more detail below). Where appropriate, the review questions were refined once the evidence had been searched and, where necessary, sub-questions were generated. The final list of review questions can be found in Appendix 8.

For questions about interventions, the PICO (Population, Intervention, Comparison and Outcome) framework was used to structure each question (see Table 2).

Table 2: Features of a well-formulated question on the effectiveness of an intervention – PICO

Population:	Which population of service users are we interested in? How can they be best described? Are there subgroups that need to be considered?
Intervention:	Which intervention, treatment or approach should be used?
Comparison:	What is/are the main alternative/s to compare with the intervention?
Outcome:	What is really important for the service user? Which outcomes should be considered: intermediate or short-term measures; mortality; morbidity and treatment complications; rates of relapse; late morbidity and readmission; return to work, physical and social functioning and other measures such as quality of life; general health status?

Questions relating to diagnosis or case identification do not involve an intervention designed to treat a particular condition, and therefore the PICO framework was not used. Rather, the questions were designed to pick up key issues specifically relevant to clinical utility, for example their accuracy, reliability, safety and acceptability to the service user.

Where review questions about service user experience were specified in the scope, the SPICE (Setting, Perspective, Intervention, Comparison and Evaluation) format was used to structure the questions (Table 3).

Table 3: Features of a well-formulated question about the experience of care (qualitative evidence) – SPICE

Setting	Where? In what context?
Perspective	For who?
Intervention (phenomenon of interest)	Which intervention/interest should be included?
Comparison	What?
Evaluation	How well? What result?
Adapted from Booth (2003).	

For each topic, addressed by one or more review questions, a review protocol was drafted by the technical team and finalised by the GDG. All protocols are included in Appendix 9.

To help facilitate the literature review, a note was made of the best study design type to answer each question. There are four main types of review question of relevance to NICE guidelines. These are listed in Table 4. For each type of question, the best primary study design varies, where 'best' is interpreted as 'least likely to give misleading answers to the question'. For questions about the effectiveness of interventions, where randomised controlled trials (RCTs) were not available, the review of other types of evidence was pursued only if there was reason to believe that it would help the GDG to formulate a recommendation.

However, in all cases, a well-conducted systematic review (of the appropriate type of study) is likely to yield a better answer than a single study.

Table 4: Best study design to answer each type of question

Type of question	Best primary study design
Effectiveness or other impact of an intervention	RCT; other studies that may be considered in the absence of RCTs are the following: internally/externally controlled before and after trial, interrupted time-series
Accuracy of information (for example, risk factor, test, prediction rule)	Comparing the information against a valid gold standard in an RCT or inception cohort study
Rates (of disease, service user experience, rare side effects)	Prospective cohort, registry, cross-sectional study
Experience of care	Qualitative research (for example, grounded theory, ethnographic research)

3.5 CLINICAL REVIEW METHODS

The aim of the clinical literature review was to systematically identify and synthesise relevant evidence from the literature in order to answer the specific review questions developed by the GDG. Thus, clinical practice recommendations are evidence-based, where possible, and, if evidence is not available, informal consensus methods are used to try and reach general agreement between GDG members (see Section3.5.7) and the need for future research is specified.

3.5.1 The search process

Scoping searches

A broad preliminary search of the literature was undertaken in March 2013 to obtain an overview of the issues likely to be covered by the scope, and to help define key areas. Searches were restricted to clinical guidelines, Health Technology Assessment (HTA) reports, key systematic reviews and RCTs. A list of databases and websites searched can be found in Appendix 10.

Systematic literature searches

After the scope was finalised, a systematic search strategy was developed to locate as much relevant evidence as possible. The balance between sensitivity (the power to identify all studies on a particular topic) and specificity (the ability to exclude irrelevant studies from the results) was carefully considered, and a decision made to utilise a broad approach to searching to maximise retrieval of evidence to all parts of the guideline. Searches were restricted to certain study designs if specified in the review protocol, and conducted in the following databases:

- Cochrane Database of Abstracts of Reviews of Effects (DARE)
- Cochrane Database of Systematic Reviews (CDSR)
- Cochrane Database of RCTs and other controlled trials (CENTRAL)
- Embase (Excerpta Medica Database)
- Health Management Information Consortium (HMIC)
- HTA database
- Medical Literature Analysis and Retrieval System Online (MEDLINE/MEDLINE In-Process)
- Psychological Information Database (PsycINFO).

The search strategies were initially developed for MEDLINE before being translated for use in other databases/interfaces. Strategies were built up through a number of trial searches and discussions of the results of the searches with the review team and GDG to ensure that all possible relevant search terms were covered. In order to assure comprehensive coverage, search terms for antenatal and postnatal mental health were kept purposely broad to help counter dissimilarities in database indexing practices and thesaurus terms, and imprecise reporting of study populations by authors in the titles and abstracts of records. The search terms for each search are set out in full in Appendix 10.

Reference management

Citations from each search were downloaded into reference management software and duplicates removed. Records were then screened against the eligibility criteria of the reviews before being appraised for methodological quality (see below). The unfiltered search results were saved and retained for future potential re-analysis to help keep the process both replicable and transparent.

Search filters

To aid retrieval of relevant and sound studies, filters were used to limit a number of searches to systematic reviews, randomized controlled trials, qualitative studies, surveys and observational studies. The search filters for systematic reviews and randomized controlled trials are adaptations of filters designed by McMaster University, Ontario, Canada. The qualitative study, surveys and observational study filter were developed in-house. Each filter comprises index terms relating to the study type(s) and associated text words for the methodological description of the design(s).

Date and language restrictions

Systematic database searches were initially conducted in April 2013 up to the most recent searchable date. Search updates were generated on a 6-monthly basis, with the final re-runs carried out in April 2014 ahead of the guideline consultation. After this point, studies were only included if they were judged by the GDG to be exceptional (for example, if the evidence was likely to change a recommendation).

Although no language restrictions were applied at the searching stage, foreign language papers were not requested or reviewed, unless they were of particular importance to a review question.

Date restrictions were not applied, except for update searches which were limited to the date of the last search conducted for NICE Clinical guideline 45. In addition searches for qualitative studies and surveys were limited to the last 15 years as service user's experiences of care pre-2000 were considered to be less relevant to the current clinical context.

Other search methods

Other search methods involved: (a) scanning the reference lists of all eligible publications (systematic reviews, stakeholder evidence and included studies) for more published reports and citations of unpublished research; (b) checking the tables of contents of key journals for studies that might have been missed by the database and reference list searches; (c) contacting included study authors for unpublished or incomplete datasets (see Appendix 5). Searches conducted for existing NICE guidelines were updated where necessary. Other relevant guidelines were assessed for quality using the AGREE instrument (AGREE Collaboration, 2003). The evidence base underlying high-quality existing guidelines was utilised and updated as appropriate.

Full details of the search strategies and filters used for the systematic review of clinical evidence are provided in Appendix 10.

Study selection and assessment of methodological quality

All primary-level studies included after the first scan of citations were acquired in full and re-evaluated for eligibility at the time they were being entered into the study information database. More specific eligibility criteria were developed for each review question and are described in the relevant clinical evidence chapters. Eligible systematic reviews and primary-level studies were critically appraised for methodological quality (risk of bias) using a checklist (see *The Guidelines Manual* [NICE, 2012a] for templates). The eligibility of each study was confirmed by at least one member of the GDG.

Unpublished evidence

Stakeholders were approached for unpublished evidence (see Appendix 4). The GDG used a number of criteria when deciding whether or not to accept unpublished data. First, the evidence must have been accompanied by a trial report containing sufficient detail to properly assess risk of bias. Second, the evidence must have been submitted with the understanding that data from the study and a summary of the study's characteristics would be published in the full guideline. Therefore, in most circumstances the GDG did not accept evidence submitted 'in confidence'. However, the GDG recognised that unpublished evidence submitted by investigators might later be retracted by those investigators if the inclusion of such data would jeopardise publication of their research. Any unpublished data used in the guideline will be specifically highlighted as such.

3.5.2 Data extraction

Quantitative analysis

Study characteristics, aspects of methodological quality, and outcome data were extracted from all eligible studies, using Review Manager 5.2 (The Cochrane Collaboration, 2012) and Excel-based forms (see Appendix 12 for study characteristics tables).

In most circumstances, for a given outcome (continuous and dichotomous), where more than 50% of the number randomised to any group were missing or incomplete, the study results were excluded from the analysis (except for the outcome 'leaving the study early', in which case, the denominator was the number randomised). Where there were limited data for a particular review, the 50% rule was not applied. In these circumstances the evidence was downgraded (see section 3.5.4).

Where possible, outcome data from an intention-to-treat analysis (ITT) (that is, a 'once-randomised-always-analyse' basis) were used. Where ITT had not been used or there were missing data, the effect size for dichotomous outcomes were

recalculated using best-case and worse-case scenarios. Where conclusions varied between scenarios, the evidence was downgraded (see section 3.5.4).

Consultation with another reviewer or members of the GDG was used to overcome difficulties with coding. Data from studies included in existing systematic reviews were extracted independently by one reviewer and cross-checked with the existing dataset. Where possible, two independent reviewers extracted data from new studies. Where double data extraction was not possible, data extracted by one reviewer was checked by the second reviewer. Disagreements were resolved through discussion. Where consensus could not be reached, a third reviewer or GDG members resolved the disagreement. Masked assessment (that is, blind to the journal from which the article comes, the authors, the institution and the magnitude of the effect) was not used since it is unclear that doing so reduces bias (Jadad et al., 1996; Berlin, 2001).

Qualitative analysis

After transcripts/reviews or primary studies of service user experience were identified (see 3.5.1), each was read and re-read and sections of the text were collected under different headings using an Excel-based form. Initially the text from the transcripts/reviews was organised using a matrix of service user experience (see Table 5).

The matrix was formed by creating a table with the eight dimensions of patient-centred care developed by the Picker Institute Europe[2], down the vertical axis, and the key points on a pathway of care (as specified by the GDG) across the horizontal axis. With regard to terminology, the GDG preferred the term 'person-centred' rather than 'patient-centred', therefore the former is used in the matrix. The Picker Institute's dimensions of patient-centred care were chosen because they are well established, comprehensive, and based on research. In addition, a variation of these dimensions has been adopted by the US Institute of Medicine (Institute of Medicine, 2001).

[2] http://www.pickereurope.org/patientcentred

Table 5: Matrix of service user experience

		Key points on the pathway of care		Themes that apply to all points on the pathway
Experience of the mental health problem				
The relationship between individual service users and professionals	Involvement in decisions and respect for preferences			
	Clear, comprehensible information and support for self-care			
	Emotional support, empathy and respect			
The way that services and systems work	Fast access to reliable health advice			
	Effective treatment delivered by trusted professionals			
	Attention to physical and environmental needs			
	Involvement of, and support for, family and carers			
	Continuity of care and smooth transitions			

Under the broad headings in the matrix, specific emergent themes were identified and coded by two researchers working independently. Overlapping themes and themes with the highest frequency count across all testimonies were extracted and regrouped using the matrix. The findings from this qualitative analysis can be found in Chapter 8.

3.5.3 Evidence synthesis

The method used to synthesize evidence depended on the review question and availability and type of evidence (see Appendix 12 for full details). Briefly, for questions about test accuracy, bivariate test accuracy meta-analysis was conducted where appropriate. For questions about the effectiveness of interventions or harms associated with interventions, standard meta-analysis or network meta-analysis was used where appropriate, otherwise narrative methods were used with clinical advice

from the GDG. In the absence of high-quality research, an informal consensus process was used (see Section 3.5.7).

3.5.4 Grading the quality of evidence

For questions about the effectiveness of interventions, the GRADE approach[3] was used to grade the quality of evidence for each outcome (Guyatt et al., 2011). For questions about the experience of care, test accuracy, and harms associated with interventions (where case-control and cohort study designs were used) methodology checklists were used to assess the risk of bias, and this information was taken into account when interpreting the evidence. The technical team produced GRADE evidence profiles (see below) using GRADEprofiler (GRADEpro) software (Version 3.6), following advice set out in the GRADE handbook (Schünemann et al., 2009). All staff doing GRADE ratings were trained, and calibration exercises were used to improve reliability (Mustafa et al., 2013).

Evidence profiles

A GRADE evidence profile was used to summarise both the quality of the evidence and the results of the evidence synthesis for each 'critical' and 'important' outcome (see Table 6 for an example of an evidence profile). The GRADE approach is based on a sequential assessment of the quality of evidence, followed by judgment about the balance between desirable and undesirable effects, and subsequent decision about the strength of a recommendation.

Within the GRADE approach to grading the quality of evidence, the following is used as a starting point:

- RCTs without important limitations provide high quality evidence
    - observational studies without special strengths or important limitations provide low quality evidence.

For each outcome, quality may be reduced depending on five factors: limitations, inconsistency, indirectness, imprecision and publication bias. For the purposes of the guideline, each factor was evaluated using criteria provided in Table 7.

For observational studies without any reasons for down-grading, the quality may be up-graded if there is a large effect, all plausible confounding would reduce the demonstrated effect (or increase the effect if no effect was observed), or there is evidence of a dose-response gradient (details would be provided under the 'other' column).

Each evidence profile includes a summary of findings: number of participants included in each group, an estimate of the magnitude of the effect, and the overall

[3] For further information about GRADE, see www.gradeworkinggroup.org

quality of the evidence for each outcome. Under the GRADE approach, the overall quality for each outcome is categorised into one of four groups (high, moderate, low, very low).

Table 6: Example of GRADE evidence profile

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other consider-ations	Intervention	Control group	Relative (95% CI)	Absolute		
Outcome 1 (measured with: any valid method; Better indicated by lower values)												
2	Random-ised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	none	47	43	-	SMD 0.20 lower (0.61 lower to 0.21 higher)	⊕⊕⊕O MODERATE	CRITICAL
Outcome 2 (measured with: any valid rating scale; Better indicated by lower values)												
4	Random-ised trials	serious[2]	no serious inconsistency	no serious indirectness	serious[1]	none	109	112	-	SMD 0.42 lower (0.69 to 0.16 lower)	⊕⊕OO LOW	CRITICAL
Outcome 3 (measured with: any valid rating scale; Better indicated by lower values)												
26	Random-ised trials	no serious risk of bias	serious[3]	no serious indirectness	no serious imprecision	none	521/ 5,597 (9.3%)	798/ 3,339 (23.9%)	RR 0.43 (0.36 to 0.51)	136 fewer per 1000 (from 117 fewer to 153 fewer)	⊕⊕⊕O MODERATE	CRITICAL
Outcome 4 (measured with: any valid rating scale; Better indicated by lower values)												
5	Random-ised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	no serious imprecision	none	503	485	-	SMD 0.34 lower (0.67 to 0.01 lower)	⊕⊕⊕⊕ HIGH	CRITICAL

[1] Optimal information size (for dichotomous outcomes, OIS=300 events; for continuous outcomes, OIS=400 participants) not met.
[2] Risk of bias across domains was generally high or unclear.
[3] There is evidence of moderate heterogeneity of study effect sizes.

Table 7: Factors that decrease quality of evidence

Factor	Description	Criteria
Limitations	Methodological quality/ risk of bias.	Serious risks across most studies (that reported a particular outcome). The evaluation of risk of bias was made for each study using NICE methodology checklists (see Section 3.5.1).
Inconsistency	Unexplained heterogeneity of results.	Moderate or greater heterogeneity (see Appendix 12 for further information about how this was evaluated)
Indirectness	How closely the outcome measures, interventions and participants match those of interest.	If the comparison was indirect, or if the question being addressed by the GDG was substantially different from the available evidence regarding the population, intervention, comparator, or an outcome.
Imprecision	Results are imprecise when studies include relatively few patients and few events and thus have wide confidence intervals around the estimate of the effect.	If either of the following two situations were met: the optimal information size (for dichotomous outcomes, OIS=300 events; for continuous outcomes, OIS=400 participants) was not achieved the 95% confidence interval around the pooled or best estimate of effect included both 1) no effect and 2) appreciable benefit or appreciable harm
Publication bias	Systematic underestimate or an overestimate of the underlying beneficial or harmful effect due to the selective publication of studies.	Evidence of selective publication. This may be detected during the search for evidence, or through statistical analysis of the available evidence.

3.5.5 Presenting evidence to the Guideline Development Group

Study characteristics tables and, where appropriate, forest plots generated with Review Manager Version 5.2 and GRADE summary of findings tables (see below) were presented to the GDG.

Where meta-analysis was not appropriate and/or possible, the reported results from each primary-level study were reported in the study characteristics table and presented to the GDG. The range of effect estimates were included in the GRADE profile, and where appropriate, described narratively.

Summary of findings tables

Summary of findings tables generated from GRADEpro were used to summarise the evidence for each outcome and the quality of that evidence (Table 8). The tables provide illustrative comparative risks, especially useful when the baseline risk varies for different groups within the population.

Table 8: Example of a GRADE summary of findings table

Patient or population: Settings: Intervention: Comparison:						
Outcomes	**Illustrative comparative risks* (95% CI)**		**Relative effect (95% CI)**	**No of Participants (studies)**	**Quality of the evidence (GRADE)**	**Comments**
	Assumed risk	**Corresponding risk**				
	Any control group	**Intervention group**				
Outcome 1 any valid rating scale		The mean outcome in the intervention group was **0.20 standard deviations lower** (0.61 lower to 0.21 higher)		90 (2 studies)	⊕⊕⊕⊖ moderate[1]	
Outcome 2 any valid rating scale		The mean outcome in the intervention group was **0.42 standard deviations lower** (0.69 to 0.16 lower)		221 (4 studies)	⊕⊕⊖⊖ low[1,2]	
Outcome 3 dichotomous data	239 per 1000	**103 per 1000** (86 to 122)	**RR 0.43** (0.36 to 0.51)	8,936 (26 studies)	⊕⊕⊕⊖ moderate[3]	
Outcome 4 any valid rating scale		The mean outcome in the intervention group was **0.34 standard deviations lower** (0.67 to 0.01 lower)		988 (5 studies)	⊕⊕⊕⊕ high	
*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).						
Note. CI=confidence interval. [1] Optimal information size (for dichotomous outcomes, OIS=300 events; for continuous outcomes, OIS=400 participants) not met. [2] Risk of bias across domains was generally high or unclear. [3] There is evidence of moderate heterogeneity of study effect sizes.						

3.5.6 Extrapolation

When answering review questions, if there is no direct evidence from a primary dataset,[4] based on the initial search for evidence, it may be appropriate to extrapolate from another data set. In this situation, the following principles were used to determine when to extrapolate:

- a primary dataset is absent, of low quality or is judged to be not relevant to the review question under consideration, and
- a review question is deemed by the GDG to be important, such that in the absence of direct evidence, other data sources should be considered, and
- non-primary data source(s) is in the view of the GDG available, which may inform the review question.

When the decision to extrapolate was made, the following principles were used to inform the choice of the non-primary dataset:

- the populations (usually in relation to the specified diagnosis or problem which characterises the population) under consideration share some common characteristic but differ in other ways, such as age, gender or in the nature of the disorder (for example, a common behavioural problem; acute versus chronic presentations of the same disorder) , and
- the interventions under consideration in the view of the GDG have one or more of the following characteristics:
    - share a common mode of action (for example, the pharmacodynamics of drug; a common psychological model of change - operant conditioning)
    - be feasible to deliver in both populations (for example, in terms of the required skills or the demands of the health care system)
    - share common side effects/harms in both populations, and
- the context or comparator involved in the evaluation of the different datasets shares some common elements which support extrapolation, and
- the outcomes involved in the evaluation of the different datasets shares some common elements which support extrapolation (for example, improved mood or a reduction in challenging behaviour).

When the choice of the non-primary dataset was made, the following principles were used to guide the application of extrapolation:

- the GDG should first consider the need for extrapolation through a review of the relevant primary dataset and be guided in these decisions by the principles for the use of extrapolation

[4] A primary data set is defined as a data set which contains evidence on the population and intervention under review

- in all areas of extrapolation datasets should be assessed against the principles for determining the choice of datasets. In general the criteria in the four principles set out above for determining the choice should be met
    - in deciding on the use of extrapolation, the GDG will have to determine if the extrapolation can be held to be reasonable, including ensuring that:
    - the reasoning behind the decision can be justified by the clinical need for a recommendation to be made
    - the absence of other more direct evidence, and by the relevance of the potential dataset to the review question can be established
    - the reasoning and the method adopted is clearly set out in the relevant section of the guideline.

3.5.7 Method used to answer a review question in the absence of appropriately designed, high-quality research

In the absence of appropriately designed, high-quality research (including indirect evidence where it would be appropriate to use extrapolation), an informal consensus process was adopted.

The process involved a member of the GDG or review team drafting a statement about what is known about the issue based on expert opinion from existing narrative reviews. The statement was circulated to the GDG and used as the basis of a group discussion.

3.5.8 Key principles for recommendations

In reviewing the evidence for mental health problems in pregnancy and/or the postnatal period the GDG were guided by the principle that much of the assessment and treatment of mental health problems in pregnancy and the postnatal period is not different from that at other times of a woman's life, and so should be guided by relevant NICE guidelines for the specific mental health problem. However, new recommendations were developed where there was new evidence specifically for this guideline:

- for an intervention that was specific to pregnancy or the postnatal period;
- that an existing recommendation needed to be clarified or modified as a result of concerns about the health of the fetus or infant;
- that changes are necessary to the context in which interventions are delivered;
    - that specific variations are necessitated by changes in a woman's mental or physical health linked to pregnancy and the postnatal period.

3.6 HEALTH ECONOMICS METHODS

The aim of the health economics was to contribute to the guideline's development by providing evidence on the cost effectiveness of interventions for women who have,

or are at risk of, mental health problems during pregnancy and the postnatal period covered in the guideline. This was achieved by:

- systematic literature review of existing economic evidence
- decision-analytic economic modelling.

Systematic reviews of economic literature were conducted in all areas covered in the guideline. Economic modelling was undertaken in areas with likely major resource implications, where the current extent of uncertainty over cost effectiveness was significant and economic analysis was expected to reduce this uncertainty, in accordance with *The Guidelines Manual* (NICE, 2012a). Prioritisation of areas for economic modelling was a joint decision between the Health Economist and the GDG. The rationale for prioritising review questions for economic modelling was set out in an economic plan agreed between NICE, the GDG, the Health Economist and the other members of the technical team. The following economic questions were selected as key issues that were addressed by economic modelling:

- cost effectiveness of formal case identification tools for depression in the postnatal period
- cost effectiveness of psychological and psychosocial interventions for the treatment of women with sub-threshold/mild to moderate depression in the postnatal period.

In addition, literature on the health-related quality of life (HRQoL) of women with mental health problems in pregnancy and postnatal period was systematically searched to identify studies reporting appropriate utility values that could be utilised in a cost-utility analysis.

The rest of this section describes the methods adopted in the systematic literature review of economic studies. Methods employed in economic modelling are described in the relevant economic sections of the evidence chapters.

3.6.1 Search strategy for economic evidence

Scoping searches

A broad preliminary search of the literature was undertaken in March 2013 to obtain an overview of the issues likely to be covered by the scope, and help define key areas. Searches were restricted to economic studies and HTA reports, and conducted in the following databases:

- Embase
- MEDLINE/MEDLINE In-Process
- HTA database (technology assessments)
- NHS Economic Evaluation Database (NHS EED).

Any relevant economic evidence arising from the clinical scoping searches was also made available to the health economist during the same period.

Systematic literature searches

After the scope was finalised, a systematic search strategy was developed to locate all the relevant evidence. The balance between sensitivity (the power to identify all studies on a particular topic) and specificity (the ability to exclude irrelevant studies from the results) was carefully considered, and a decision made to utilise a broad approach to searching to maximise retrieval of evidence to all parts of the guideline. Searches were restricted to economic studies and health technology assessment reports, and conducted in the following databases:

- Embase
- HTA database (technology assessments)
- MEDLINE/MEDLINE In-Process
- NHS EED
- PsycINFO.

Any relevant economic evidence arising from the clinical searches was also made available to the health economist during the same period.

The search strategies were initially developed for MEDLINE before being translated for use in other databases/interfaces. Strategies were built up through a number of trial searches, and discussions of the results of the searches with the review team and GDG to ensure that all possible relevant search terms were covered. In order to assure comprehensive coverage, search terms for the guideline topic were kept purposely broad to help counter dissimilarities in database indexing practices and thesaurus terms, and imprecise reporting of study populations by authors in the titles and abstracts of records.

For standard mainstream bibliographic databases (CINAHL, Embase, MEDLINE and PsycINFO) search terms for the guideline topic combined with a search filter for health economic studies. For searches generated in topic-specific databases (HTA, NHS EED) search terms for the guideline topic were used without a filter. The sensitivity of this approach was aimed at minimising the risk of overlooking relevant publications, due to potential weaknesses resulting from more focused search strategies. The search terms are set out in full in Appendix 11.

Reference Management

Citations from each search were downloaded into reference management software and duplicates removed. Records were then screened against the inclusion criteria of the reviews before being quality appraised. The unfiltered search results were saved and retained for future potential re-analysis to help keep the process both replicable and transparent.

Search filters

The search filter for health economics is an adaptation of a pre-tested strategy designed by CRD (InterTASC Information Specialists' SubGroup, 2007). The search filter is designed to retrieve records of economic evidence (including full and partial economic evaluations) from the vast amount of literature indexed to major medical databases such as MEDLINE. The filter, which comprises a combination of controlled vocabulary and free-text retrieval methods, maximises sensitivity (or recall) to ensure that as many potentially relevant records as possible are retrieved from a search. A full description of the filter is provided in Appendix 11.

Date and language restrictions

Systematic database searches were initially conducted in April 2013 up to the most recent searchable date. Search updates were generated on a 6-monthly basis, with the final re-runs carried out in April 2014 ahead of the guideline consultation. After this point, studies were included only if they were judged by the GDG to be exceptional (for example, the evidence was likely to change a recommendation).

Although no language restrictions were applied at the searching stage, foreign language papers were not requested or reviewed, unless they were of particular importance to an area under review. All new searches were restricted to research published from 1998 onwards in order to obtain data relevant to current healthcare settings and costs. All update searches were restricted to the date of the last search conducted for NICE clinical guideline 45 (NICE, 2007a).

Other search methods

Other search methods involved scanning the reference lists of all eligible publications (systematic reviews, stakeholder evidence and included studies from the economic and clinical reviews) to identify further studies for consideration.

Full details of the search strategies and filter used for the systematic review of health economic evidence are provided in Appendix 11.

3.6.2 Inclusion criteria for economic studies

The following inclusion criteria were applied to select studies identified by the economic searches for further consideration:

- Only studies from Organisation for Economic Co-operation and Development countries were included, as the aim of the review was to identify economic information transferable to the UK context.
- Only English language papers were considered.
- Studies published from 2006 onwards were included. This date restriction was imposed to obtain data relevant to current healthcare settings and costs.
- Selection criteria based on types of clinical conditions and service users as well as interventions assessed were identical to the clinical literature review.

- Studies were included provided that sufficient details regarding methods and results were available to enable the methodological quality of the study to be assessed, and provided that the study's data and results were extractable. Poster presentations, abstracts, dissertations, commentaries and discussion publications were excluded.
- Full economic evaluations that compared two or more relevant interventions and considered both costs and consequences, as well as costing analyses comparing only costs between two or more interventions, were included in the review.
- Economic studies were included if they used clinical effectiveness data from an RCT, a prospective cohort study, or a systematic review and meta-analysis of clinical studies. Studies that had a mirror-image or other retrospective design were excluded from the review. Also, studies that utilised clinical effectiveness parameters based mainly on expert opinion or assumptions were excluded from the review.
- Studies were included only if the examined interventions and populations under consideration were clearly described.

3.6.3 Applicability and quality criteria for economic studies

All economic papers eligible for inclusion were appraised for their applicability and quality using the methodology checklist for economic evaluations recommended by NICE (NICE, 2012a). The methodology checklist for economic evaluations was also applied to the economic models developed specifically for this guideline. All studies that fully or partially met the applicability and quality criteria described in the methodology checklist were considered during the guideline development process, along with the results of the economic modelling conducted specifically for this guideline. The completed methodology checklists for all economic evaluations considered in the guideline are provided in Appendix 20.

3.6.4 Presentation of economic evidence

The economic evidence considered in the guideline is provided in the respective evidence chapters, following presentation of the relevant clinical evidence. The references to included studies and the respective evidence tables with the study characteristics and results are provided in Appendix 21. Methods and results of economic modelling undertaken alongside the guideline development process are presented in the relevant evidence chapters. Characteristics and results of all economic studies considered during the guideline development process (including modelling studies conducted for this guideline) are summarised in economic evidence profiles accompanying respective GRADE clinical evidence profiles in Appendix 22.

3.6.5 Results of the systematic search of economic literature

The titles of all studies identified by the systematic search of the literature were screened for their relevance to the topic (that is, economic issues and information on HRQoL). References that were clearly not relevant were excluded first. The abstracts

of all potentially relevant studies (15 references) were then assessed against the inclusion criteria for economic evaluations by the health economist. Full texts of the studies potentially meeting the inclusion criteria (including those for which eligibility was not clear from the abstract) were obtained. Studies that did not meet the inclusion criteria, were duplicates, were secondary publications of one study, or had been updated in more recent publications were subsequently excluded. Economic evaluations eligible for inclusion (9 studies in 12 publications) were then appraised for their applicability and quality using the methodology checklist for economic evaluations. Finally, 9 economic studies that fully or partially met the applicability and quality criteria were considered at formulation of the guideline recommendations.

3.7 USING NICE EVIDENCE REVIEWS AND RECOMMENDATIONS FROM EXISTING NICE CLINICAL GUIDELINES

When review questions overlap and evidence from another guideline applies to a question in the current guideline, it might be desirable and practical to incorporate or adapt recommendations published in NICE guidelines. Adaptation refers to the process by which an existing recommendation is modified in order to facilitate its placement in a new guideline. Incorporation refers to the placement of a recommendation that was developed for another guideline into a new guideline, with no material changes to wording or structure. Incorporation would be used in relatively rare circumstances, as cross-referring to the other guideline will often be all that is necessary.

Incorporation or adaptation is likely to be substantially more complex where health economics were a major part of the decision making. In these circumstances, these methods are only used rarely after full and detailed consideration.

3.7.1 Incorporation

In the current guideline, the following criteria were used to determine when a recommendation could be incorporated:

- a review question in the current guideline was addressed in another NICE guideline
- evidence for the review question and related recommendation(s) has not changed in important ways
- evidence for the previous question is judged by the GDG to support the existing recommendation(s), and be relevant to the current question
- the relevant recommendation can 'stand alone' and does not need other recommendations from the original guideline to be relevant or understood within the current guideline.

3.7.2 Adaptation

The following criteria were used to determine when a recommendation could be adapted:

- a review question in the current guideline is similar to a question addressed in another NICE guideline
- evidence for the review question and related recommendations has not changed in important ways
- evidence for the previous question is judged by the GDG to support the existing recommendation(s), and be relevant to the current question
- the relevant recommendation can 'stand alone' and does not need other recommendations from the original guideline to be relevant
- contextual evidence, such as background information about how an intervention is provided in the healthcare settings that are the focus of the guideline, informs the re-drafting or re-structuring of the recommendation but does not alter its meaning or intent (if meaning or intent were altered, a new recommendation should be developed).

In deciding whether to choose between incorporation or adaption of existing guideline recommendations, the GDG considered whether the direct evidence obtained from the current guideline dataset was of sufficient quality to allow development of recommendations. It was only where (a) such evidence was not available or insufficient to draw robust conclusions and (b) where methods used in other NICE guidelines were sufficiently robust that the 'incorporate and adapt' method could be used. Recommendations were only incorporated or adapted after the GDG had reviewed evidence supporting previous recommendations and confirmed that they agreed with the original recommendations.

When adaptation is used, the meaning and intent of the original recommendation is preserved but the wording and structure of the recommendation may change. Preservation of the original meaning (that is, that the recommendation faithfully represents the assessment and interpretation of the evidence contained in the original guideline evidence reviews) and intent (that is, the intended action[s] specified in the original recommendation will be achieved) is an essential element of the process of adaptation.

3.7.3 Roles and responsibilities

The guideline review team, in consultation with the guideline Facilitator and Chair, were responsible for identifying overlapping questions and deciding if it would be appropriate to incorporate or to adapt following the principles above. For adapted recommendations, at least two members of the GDG for the original guideline were consulted to ensure the meaning and intent of the original recommendation was preserved. The GDG confirmed the process had been followed, that there was

insufficient evidence to make new recommendations, and agreed all adaptations to existing recommendations.

In evidence chapters where incorporation and adaptation have been used, the original review questions are listed with the rationale for the judgement on the similarity of questions. Tables are then provided that set out the original recommendation, a brief summary of the original evidence, the new recommendation, and the reasons for adaptation. For an adapted recommendation, details of any contextual information are provided, along with information about how the GDG ensured that the meaning and intent of the adapted recommendation was preserved.

3.7.4 Drafting of adapted recommendations

The drafting of adapted recommendations conformed to standard NICE procedures for the drafting of guideline recommendations, preserved the original meaning and intent, and aimed to minimise the degree or re-writing and re-structuring.

3.8 LINKING EVIDENCE TO RECOMMENDATIONS

Once the clinical and health economic evidence was summarised, the GDG drafted the recommendations. In making recommendations, the GDG took into account the trade-off between the benefits and harms of the intervention/instrument, as well as other important factors, such as economic considerations, values of the GDG and society, the requirements to prevent discrimination and to promote equality[5], and the GDG's awareness of practical issues (Eccles et al., 1998; NICE, 2012a).

Finally, to show clearly how the GDG moved from the evidence to the recommendations, each chapter has a section called 'linking evidence to recommendations'. Underpinning this section is the concept of the 'strength' of a recommendation (Schünemann et al., 2003). This takes into account the quality of the evidence but is conceptually different. Some recommendations are 'strong' in that the GDG believes that the vast majority of healthcare professionals and service users would choose a particular intervention if they considered the evidence in the same way that the GDG has. This is generally the case if the benefits clearly outweigh the harms for most people and the intervention is likely to be cost effective. However, there is often a closer balance between benefits and harms, and some service users would not choose an intervention whereas others would. This may happen, for example, if some service users are particularly averse to some side effect and others are not. In these circumstances the recommendation is generally weaker, although it may be possible to make stronger recommendations about specific groups of service users. The strength of each recommendation is reflected in the wording of the recommendation, rather than by using ratings, labels or symbols.

[5]See NICE's equality scheme: www.nice.org.uk/aboutnice/howwework/NICEEqualityScheme.jsp

Where the GDG identified areas in which there are uncertainties or where robust evidence was lacking, they developed research recommendations. Those that were identified as 'high priority' were developed further in the NICE version of the guideline, and presented in Appendix 15.

3.9 STAKEHOLDER CONTRIBUTIONS

Professionals, service users, and companies have contributed to and commented on the guideline at key stages in its development. Stakeholders for this guideline include:

- service user and carer stakeholders: national service user and carer organisations that represent the interests of people whose care will be covered by the guideline
- local service user and carer organisations: but only if there is no relevant national organisation
- professional stakeholders' national organisations: that represent the healthcare professionals who provide the services described in the guideline
- commercial stakeholders: companies that manufacture drugs or devices used in treatment of the condition covered by the guideline and whose interests may be significantly affected by the guideline
- providers and commissioners of health services in England and Wales
- statutory organisations: including the Department of Health, the Welsh Assembly Government, NHS Quality Improvement Scotland, the Care Quality Commission and the National Patient Safety Agency
- research organisations: that have carried out nationally recognised research in the area.

NICE clinical guidelines are produced for the NHS in England and Wales, so a 'national' organisation is defined as one that represents England and/or Wales, or has a commercial interest in England and/or Wales.

Stakeholders have been involved in the guideline's development at the following points:

- commenting on the initial scope of the guideline and attending a scoping workshop held by NICE
- contributing possible review questions and lists of evidence to the GDG
- commenting on the draft of the guideline.

3.10 VALIDATION OF THE GUIDELINE

Registered stakeholders had an opportunity to comment on the draft guideline, which was posted on the NICE website during the consultation period. Following the consultation, all comments from stakeholders and experts (see Appendix 7) were

responded to, and the guideline updated as appropriate. NICE also reviewed the guideline and checked that stakeholders' comments had been addressed.

Following the consultation period, the GDG finalised the recommendations and the NCCMH produced the final documents. These were then submitted to NICE for a quality assurance check. Any errors were corrected by the NCCMH, then the guideline was formally approved by NICE and issued as guidance to the NHS in England and Wales.

4 THE ORGANISATION OF PERINATAL MENTAL SERVICES

The 2007 guideline review of the organisation of perinatal mental health services has not been updated because it was outside the remit for this update. There have been slight amendments to the language used in the recommendations so that they are consistent with the updated recommendations in this guideline, but there have been no significant changes to the context and meaning of the recommendations.

In addition, one recommendation (4.5.1.5) that was previously located in the chapter 'The prediction and detection of mental illness during pregnancy and the postnatal period' in the 2007 guideline[6] has been moved to this chapter because it is related to the work of perinatal mental health services, which is the focus of this chapter. The review itself has not been updated.

The 2007 review is summarised below (see Appendix 27 for the full 2007 guideline chapter, to provide context for the recommendations that were made in 2007 and that still stand for this guideline). However, it is important to bear in mind that no new evidence has been reviewed or added to this chapter.

4.1 THE 2007 STRUCTURE OF SERVICES

The 2007 guideline took as its starting point a review of the structure of services based on two surveys commissioned by the GDG.

4.1.1 2007 Survey of primary care

In order to inform the 2007 guideline development process, that GDG commissioned a survey of the perinatal mental health services within primary care. The purpose of the survey was to investigate the structure of mental health services for pregnant and postnatal women throughout England and Wales. See Appendix 25 for a copy of the survey and Appendix 26 for full results of the survey. The results of that survey suggested that (in 2007) specialist primary care provision for women with mental health problems during pregnancy and the postnatal period was patchy. It was estimated that under the old NHS system of primary care trusts (PCTs), only around 25% of PCTs had a fully developed and implemented policy for antenatal and postnatal mental health.

4.1.2 2007 Survey of specialist perinatal services

A survey was also conducted of all potential provider trusts of specialist mental health services for pregnant and postnatal women in England and Wales and a similar picture of unequal distribution of services was suggested. To summarise the

[6] 'The prediction and detection of mental illness during pregnancy and the postnatal period' chapter from the 2007 guideline has largely been replaced by chapter 5 ('Case identification and assessment') in this guideline.

results of the survey of specialist perinatal services, very patchy provision was found, with the expertise concentrated in one or two areas. The distribution of services and their precise location also varied considerably. See the full 2007 chapter in Appendix 27 for further information.

4.2 ESTIMATING THE NEED FOR SERVICES

Service functions and the structures to ensure their effective delivery should be based on an understanding of the nature of mental health problems and their epidemiology, which are summarised in Chapter 2. The number of live births in 2004 in England and Wales was 639,721 (Office for National Statistics, 2006), which is an average of 13 per 1000, although the rate will vary considerably from area to area. A GP with an average-sized list (1,800 patients) may therefore expect somewhere between 15 and 27 live births on his or her list each year.

4.2.1 Common mental health disorders during pregnancy and the postnatal period

The epidemiology of perinatal disorder has been covered in Chapter 2; it is briefly considered again here, to give an indication of the likely need for services. As is apparent from Chapter 2, the epidemiology of antenatal and postnatal mental health problems is not well understood and caution must be exercised in basing service structures on this data. Careful and critical analysis of this and other locally collected data must be used when developing local services.

Common mental health problems during pregnancy and the postnatal period include depression and anxiety disorders, such as panic disorder, OCD and PTSD. An estimated 10% to 15% of women suffer from depression after the birth of an infant (Brockington, 1996; Nonacs & Cohen, 1998); in England and Wales this is between 64,000 and 94,000 women a year and is equivalent to between two and three women per year on the average GP list and 100 to 150 per 1000 live births. Prevalence data for anxiety disorders during the perinatal period are not as reliable. The Office for National Statistics estimates that the prevalence of anxiety is around 4% of men and 5% of women (Office for National Statistics, 2006). This would mean that around 30,000 women giving birth per year are also likely to be suffering from anxiety, with two or three women per year on the average GP list (50 per 1000 live births). A key role of maternity and primary care services in antenatal and postnatal mental healthcare is the identification of mental health problems. Case identification of mental health problems in pregnancy and the postnatal period is covered in Chapter 5.

It has been estimated that 50% of people with depression (that is, all those with depression, not just those with depression occurring in the postnatal period) are not identified (Williams et al., 1995). This means that around half of the 128 to 192 pregnant or postnatal women who develop depression per 100,000 population may present to primary care mental health services each year (that is, 50 to 75 per

1000 live births). A similar or lower figure might reasonably be expected for anxiety disorders, with fewer disorders being identified than for depression.

For the vast majority of these women, professional help will be provided solely by primary healthcare services. However, this is not always the case; for example, around 3% to 5% of women giving birth have moderate or severe depression, with about 1.7% being referred to specialist mental health services (Cox et al., 1993; O'Hara & Swain, 1996). Thus, around 17 women per 1000 live births would be referred to specialist mental health services with depression postnatally. Again, it is reasonable to expect the figures for anxiety disorders to follow the national trend, with a lower rate of referral through to specialist services.

4.2.2 Severe mental illness during pregnancy and the postnatal period

First presentations of severe mental illness, primarily schizophrenia and bipolar disorder, in the perinatal period are rare, with a rate in the region of two per thousand resulting in hospital admissions (based on admission as a proxy for psychosis) (Kendell et al., 1987). These episodes are associated with a clustering of admissions in the first month after the birth (1 per 2,000 live births). More common, particularly with bipolar disorder, is the exacerbation of an existing disorder, with some studies reporting relapse rates for bipolar disorder approaching 50% in the antenatal period and 70% in the postnatal period (Viguera et al., 2000). These women, along with others suffering from severe depression and other severe disorders such as severe anxiety disorders or personality disorders, will benefit from referral to specialist mental health services.

These figures, along with data obtained from a survey in the Nottingham area (Oates, 2000), give some indication of the range of presentations to specialist services, with estimates of the number of new presentations in the range of 18 to 30 per 100,000 head of population and a further 12 to 24 per 100,000 presentations of already identified disorder, giving a total estimate in the region of 30 to 54 per 100,000.

Some of these women will require inpatient care. These include those with postpartum psychosis and a number of women with severe depressive disorders. Some of these are cared for in mother and baby units (MBUs). A survey, as part of a larger study of alternatives to admission in the UK, identified 19 units: MBUs and mother and baby facilities (hospitals where one or two mother and baby beds are provided in the absence of a designated unit) with 126 available beds (personal communication, S. Johnson, 30 June 2006).

Determining the need for specialist services, including where appropriate specialist perinatal teams and the number of inpatient facilities, their size and location, is difficult for a number of reasons. Firstly, the incidence of severe mental illness requiring inpatient care varies across the country, with much higher morbidity in inner city areas compared with suburban or rural areas. Secondly, the local structure of services (for example, the presence of crisis and home treatment teams) may also

impact significantly on the use of inpatient services (Killaspy et al., 2006). Thirdly, the presence of specialist perinatal services that have responsibility for the coordination/delivery of care to women with severe perinatal psychiatric disorders, and the way in which they are designed, may also impact on referral rates and on bed usage. There is also some evidence to suggest that the provision of specialist inpatient services without specialist community services to coordinate such care can be associated with higher inpatient bed usage. Fourthly, significant numbers of MBUs also use a number of their beds for parenting assessments; that is, the assessment of a woman's capacity to care for her child. These assessments, which can be extended over several weeks, may occupy up to 80% of beds in some MBUs and as such may limit the capacity of the units to care effectively for emergency admissions.

In arriving at estimates of need for inpatient services, the balance of geographical proximity and the need to develop economies of scale also need to be taken into account. In addition, caution is required when determining bed requirements from average bed-use data; there will be considerable variation in demand for beds and duration of use, which can seriously undermine calculations based on averages (Gallivan et al., 2002). The 2007 guideline review suggested that, given the 2007 provision of approximately 110 specialist beds, between 30 and 50 additional perinatal specialist beds would be required to meet the needs of women with severe mental illness who required admission in the perinatal period. This assumed that all units would be equally accessible but, given the geography and population distribution of England and Wales, it is likely that additional beds would be required to provide reasonable access and to provide the capacity to respond appropriately to emergency admissions. This suggested that between 60 and 80 additional beds would be required.

4.3 THE FUNCTIONS OF SERVICES FOR WOMEN, THEIR PARTNERS AND CARERS IN PREGNANCY AND THE POSTNATAL PERIOD

When identifying the key functions of any healthcare system, the needs of the patient are central. Anyone with a mental health problem, regardless of other factors, should have:

- the disorder detected effectively
- effective assessment and referral to appropriate services when necessary
- timely, appropriate treatment
- accurate information about the disorder and the benefits and risks associated with treatment, including psychotropic medication
- provision of care in the most appropriate setting
- appropriate communication about their care, with other services as required and without unnecessary breaches of confidentiality or stigmatising procedures
- choice.

For women with mental health problems during pregnancy and postnatally, the clinical context is complicated by the needs of the fetus and infant, such as the safety of drugs during pregnancy and breastfeeding, and by the woman's psychological adjustment to pregnancy, motherhood or having an additional child while experiencing mental illness. Services also need to take into account the needs of fathers/partners, carers and other children in the family. Therefore, services need to be tailored to meet these needs, which may include the provision of specialist inpatient services, integration of specific mental health services and maternity services, and dedicated treatment programmes. These must be provided in a timely fashion to ensure that treatments giving relief to the woman do so before her condition has damaged the health and development of the fetus and other family members. This is particularly relevant for the provision of psychological treatment. Such services may be configured in different ways to provide the same functions to patients, dependent on local considerations, such as population density and variations in morbidity.

In meeting the mental health needs of women in the perinatal period, services should seek to provide the most effective and accessible treatments in the least intrusive and disruptive manner. This principle, of stepped care, has helped organise services in other aspects of mental health provision (for example, NICE [2004a]). Professionals, from core primary care teams such as health visitors and GPs through to perinatal psychiatrists, and women and families themselves, are all involved in delivering an effective mental health service for women in pregnancy and the postnatal period. A key function is the development and implementation of clear care pathways and effective working between different professionals that always hold the women (and fetus/infant) at the centre of consideration.

In general, early steps in the pathway will be provided by generalist primary care professionals and generalist maternity services. The model includes mental health professionals such as counsellors and primary care mental health workers as appropriate. When there is a requirement for more intensive treatments, more specialist professionals will need to be involved. Some women (and their fetus/infant) may need the intervention of a specialist inpatient setting. Specialist perinatal teams may provide input (including advice and consultations, as well as direct care) at a variety of points in an individual woman's care pathway.

4.3.1 General healthcare services (including primary care and maternity services)

All pregnant women have contact with general healthcare services. Maternity services may be a mix of community services, which may be midwife-led, and hospital-based services, including hospital-based midwives and obstetricians. It is these professionals who are well placed to identify women with a history of, or current, mental health problem in pregnancy. The case identification of mental health problems in pregnancy (and the postnatal period) is covered in Chapter 5.

4.3.2 Primary care mental health services

The vast majority of women with mental health problems during the perinatal period present to, and are treated solely by, primary care services. Primary care mental health services include GPs, practice counsellors and psychological therapists, practice nurses, health visitors, midwives and primary care mental health workers. Key functions of these services are to: provide assessment, treatment and care as necessary; liaise with and make appropriate referrals to specialist services; make appropriate use of service user support groups; identify risk, including risk to the infant's health and wellbeing, or that of other children in the family; and communicate with other services.

4.3.3 Specialist mental health services including specialist perinatal mental health services

Women requiring specialist care may be treated by general mental health services, specialist perinatal mental health services, or combinations of these services. The functions of specialist mental health services, including specialist perinatal services, are as follows:

- assessment of women with moderate and severe mental health problems (or those with milder but treatment-resistant disorder) during pregnancy and the postnatal period, including assessment of the risk of relapse of existing disorder during pregnancy, childbirth or the postnatal period
- treatment of mental health problem during pregnancy and the postnatal period
- provision of intensive services, such as crisis, home treatment and inpatient services and, in the case of some specialist perinatal services, the provision of specialist inpatient beds
- communication with primary care, maternity and obstetric services and, where appropriate, coordination and management of care pathways and service access
- provision of specialist consultation and advice to services providing treatment and care to patients with existing disorder who are planning a pregnancy or who become pregnant, and to services managing women with less severe disorders; this may include advice on care, treatment, mother-infant relationships, child protection issues and diagnosis
- liaison with primary care and maternity services concerning the care of women with moderate to severe mental health problems
- education and training for maternity and primary and secondary care mental health services.

4.3.4 Inpatient services

Women presenting to secondary care mental health services during pregnancy or the postnatal period may require inpatient care. MBUs are designed to address a number of challenges, including the need for specialist expertise in the treatment of

severe perinatal illness, the need to support the development of the mother-infant relationship through a joint admission, and the provision of an environment that is safe and appropriate to the care of a young infant (for example, the presence of specialist nursery nurses and the avoidance of the severe disturbance seen on many general inpatient wards) and to the physical needs of pregnant and postnatal women. The functions of inpatient services for women with mental health problems during pregnancy and the postnatal period include:

- assessment of mental illness, including risk assessment and assessment of ability to care for the infant
- provision of expert care of women requiring admission
- in MBUs, the expert provision of safe care for the infants of women admitted
- support for the woman in caring for and developing a relationship with her baby, wherever appropriate fostering the involvement of the partner or other carers
- liaison and integrated working with other services, including maternity and obstetric services, GPs, and maternity-based and community mental health services.

A key factor in the decision to admit a woman with her infant is consideration of the welfare of the infant. That is, whether it is better for the infant to stay with his or her mother or whether he or she should be cared for by another family member while the woman receives inpatient treatment. In 2007, where specialist units were available, women were usually admitted with their infants unless there was good reason not to, for example, the woman preferring not to have her child with her or the child requiring specialist medical care not available in the unit. However, admission to a unit was also influenced by geographical proximity (Brockington, 1996).

There are few formal evaluations of the provision of MBUs and fewer still of the cost effectiveness of this model of care provision. A systematic search of the literature for the 2007 guideline identified no economic studies of inpatient units or specialist perinatal teams, and only one study that assessed the cost effectiveness of a specialised psychiatric day-hospital unit for the treatment of women with depression in the postnatal period was found (Boath et al., 2003) (see Appendix 24). In this study, the economic analysis was conducted alongside a prospective cohort study carried out in the UK. The study population consisted of 60 women with an EPDS score >12 and a diagnosis of major or minor depressive disorder (according to research diagnostic criteria [RDC]), who had an infant aged between 6 weeks and 1 year. The comparator of the analysis was a neighbouring area providing routine primary care by GPs and health visitors with referrals into secondary care. The primary clinical outcome used in the economic analysis was the number of women successfully treated, defined as no longer fulfilling RDC for major or minor depressive disorder. The analysis adopted a societal perspective and costs and outcomes were measured over a period of 6 months. The analysis demonstrated that

the day-hospital unit resulted in a significantly higher number of women successfully treated compared with routine primary care, but at an additional cost of £1,945 per successfully treated woman (1992/93 prices). The cost per successfully treated woman in the routine primary care group was estimated at £2,710. Since the NHS was prepared to pay £2,710 for a successful outcome achieved in routine primary care, the authors concluded that the unit was a cost-effective alternative treatment approach, providing additional benefit at an incremental cost below what the NHS was already paying for the treatment of women with depression in the postnatal period. However, the study had a number of limitations, such as the cohort design, which was subject to systematic bias and confounding variables, the short time horizon of the analysis and, most importantly, the selection of the comparator (that is, non-specialised primary care with only occasional referrals to specialists), which may have led to overestimation of incremental benefits associated with the unit.

4.4 MANAGED CLINICAL NETWORKS

4.4.1 Introduction to managed clinical networks

The 2007 guideline review concluded that since the precise structure of services will vary in different parts of the country based on local factors, including the organisation of existing mental health services, the demographic profile of the local population and geographical issues, the provision of services needs to be seen in terms of standard features that can be adopted by any service and adapted to meet local need in order to deliver integrated care. The 2007 guideline conceptualised this using a managed network model, with managed clinical networks defined as linked groups of health professionals and organisations from primary, secondary and tertiary care working in a coordinated manner, unconstrained by existing professional and service boundaries, to ensure equitable provision of high-quality clinically effective services.

Models of managed clinical networks

A number of models for the development of managed clinical networks have been developed and these have been reviewed by Goodwin and colleagues (2004). Goodwin describes three broad types of network: enclave, hierarchical and individualistic. All three have potential benefits and no one model is held to be superior to the others. In fact, in practice most networks have elements of all three models. However, in view of the potential functions of a perinatal mental health network, the hierarchical model was judged as the most appropriate by the 2007 GDG. This model is defined as having 'an organisational core and authority to regulate the work of members via joint provision, inspection and/or accreditation'. Such networks are held to be most successful in coordinating and controlling a pre-defined task that involves complex division of labour, and therefore were seen as the most appropriate structure for a perinatal mental health network, where agreement on care pathways, thresholds for admission and allocation of resources to community and inpatient services will need to be determined. In contrast to some networks based on this model, for example cancer networks, the limitations of the

2007 evidence base suggested that the emphasis in a perinatal network would be on joint provision and ensuring the quality of services, as the evidence base was insufficient to develop accreditation systems.

Goodwin and colleagues (2004) also described the characteristics of successful networks and these included:

- Central coordination – key for hierarchical networks and should be financed, proactive and with the possibility of a 'neutral manager or agency' where there are competing interests.
- Clear mission statement and unambiguous rules of engagement.
- Inclusivity – ensuring all agencies and individuals gain ownership of the network.
- Manageable size – large networks should be avoided due to high administrative costs and the inertia that can develop.
- Cohesion – strategies should be developed aimed at achieving network cohesion, which could include joint finance arrangements, pooled budgets, agreed care protocols and common targets. A 'boundary spanner', acting as an intermediary between organisations and agencies, allows individualistic networks to function effectively and helps hierarchical networks engage with peripheral agencies. It can be a key enabler in promoting network cohesion across all network types.
- Ownership facilitated by formalised contracts and agreements, with avoidance of over-regulation.
- Leadership – respected professional leaders who will promote the network to peers should be actively engaged.
- Avoidance of network domination by a professional elite or a particular organisational culture.
- Response to the needs of network members in such a way that the network remains relevant and worthwhile.
- Professionals in networks providing the mandate to allow managers to manage and govern their activities.

Such models have been adopted in the UK for the development of a number of medical services, including those for cancer (34 cancer networks were developed in 2001 in England), cardiovascular care, emergency care and genitourinary medicine. In addition, they have been extensively promoted in the Scottish healthcare system.

Developing a perinatal mental health managed network

A central concern in developing a perinatal mental health managed network was ensuring that women with mental health problems during pregnancy and the postnatal period have appropriate access to both specialist perinatal expertise and, where necessary, inpatient care. This factor is important in determining the size of a network with coordinated inpatient services. Such units and the networks that are built around them would need to be in accordance with the factors associated with success identified by Goodwin and colleagues (2004), be clinically and economically

viable and be geographically located so that undue burdens are not placed on patients and their families in accessing them.

Adopting a hierarchical model for a perinatal network would require that the network has:

- an identified manager with clearly specified and delegated responsibilities, who may be independent of any one element of the network or located in the element of the network that contains the inpatient unit(s) and has responsibilities to ensure that the relationship within the network is properly developed and maintained
- a clear mission statement – in which the expectations of all parties are clearly set out
- a system – normally a management board that recognises and guarantees the ownership of the network by all agencies, including clinicians, commissioners and managers, and supports the development of a shared and reflective network culture
- a size that delivers appropriate economies of scale but which does not generate high administrative costs and inertia
- clearly specified and contracted finance arrangements, agreed referral and care protocols and information systems to support the effective operation of the network
- active professional leadership and full multidisciplinary involvement.

Advantages of perinatal mental health managed networks

Perinatal mental health managed networks were judged by the 2007 GDG to confer a number of potential advantages including:

- the effective concentration of expertise and the identification of dedicated time and explicit responsibility for the delivery of appropriate care to mentally ill women and their families
- the identification of clear care pathways, a threshold for referrals and evidence-based protocols to support healthcare professionals in identifying and managing the most serious disorders presenting around childbirth
- improved liaison with, and effective monitoring and support of, maternity services
- providing more widely available up-to-date information about the impact of psychotropic medication in pregnancy and breastfeeding and advice on how to assess and effectively communicate the risks and benefits of their use in an individual woman
- playing a key role in training, education and raising awareness for maternity services, general mental health services and primary care to enable non-specialists to be as effective and confident about perinatal mental health as possible and have access to advice about where their limits lie
- more equitable and cost-effective use of inpatient services, with more effective evaluation of the likely risks and benefits of admission for particular women and the purpose of admission to an MBU

- more favourable outcomes in terms of reduced mortality and morbidity, and increased patient satisfaction
- more timely services for those women who need treatment for their mental health problems urgently because their illnesses may have a disproportionate effect on the fetus
- improved access to psychological therapies
- supporting standard setting and monitoring, participation in research and the integration of learning from national schemes such as the Confidential Enquiry into Maternal and Child Health (CEMACH).

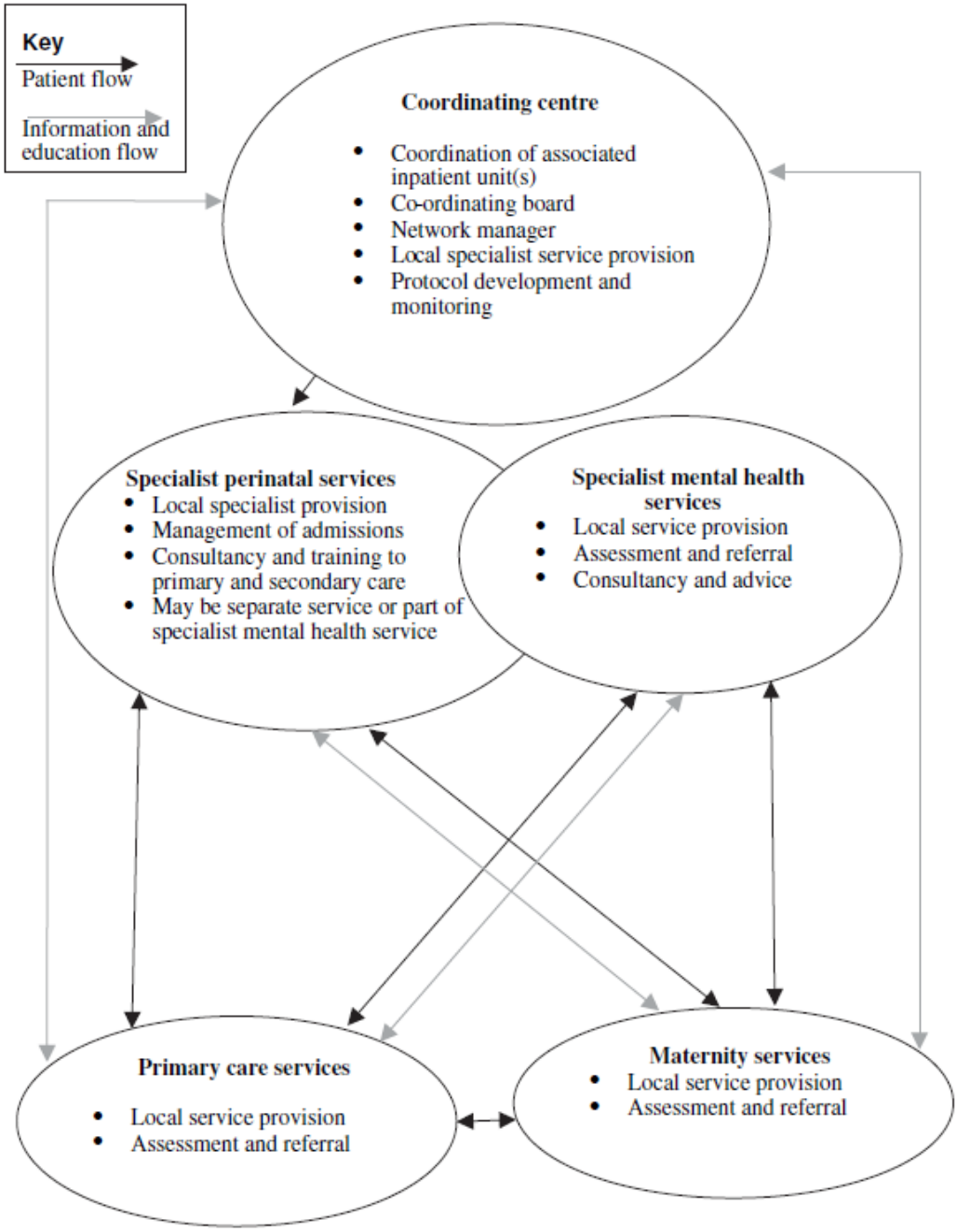
Structure of perinatal mental health managed networks

An outline of a perinatal clinical network model is set out in the 2007 guideline and replicated in Figure 2. This model, in line with a stepped-care approach, assumes that inpatient care in a network could be provided on behalf of the network by one or more member organisations, depending on the identified need in the network and its geographical structure.

In the model outlined, the managed network would be coordinated by a network board, with a core coordinating team drawn from senior staff in relevant specialist perinatal teams, maternity services, secondary care mental health services, and primary care, as well as commissioners and service user and carer representatives. The board would have responsibility for overseeing the development of protocols and pathways for the coordination of care between services, implementing good practice, coordinating expert clinical advice, management and local strategy. The aim was to ensure that services work together to improve quality of care and address any inequalities in provision and access in the area covered by the network.

The precise area covered by each network would be determined by local need, but one determinant would be the need for effective use of inpatient services. In determining the need for inpatient beds, a number of factors need to be considered; these include the critical mass of expertise to ensure effective treatment of women and their infants and the trade-off of geographical proximity. Units of fewer than eight to ten beds may be less cost effective, and units of fewer than four to six beds may not be able to maintain sufficient staffing and expertise to be able to respond comprehensively to the needs of women and their infants; units above 12 beds are likely to present complex organisational and management problems.

Figure 2: Perinatal clinical network



In this model, local specialist perinatal services have a key role in linking specialist inpatient services with general mental health, maternity and primary care services. Such specialist services would vary in size and composition according to local circumstances. They may include 'stand-alone' specialist perinatal services providing a broad community-based service, services linked to liaison psychiatry or liaison obstetric services, or services linked to community mental health services. Indeed, given local variations in morbidity and service structures, the latter models may be the most effective way to provide services in some areas rather than stand-alone specialist perinatal mental health teams given that there is no direct evidence for the effectiveness of such teams within the UK healthcare system. Also, there is

patchy evidence for the effectiveness of other functional mental health teams in the NHS, including crisis teams, assertive outreach teams (for example, Killaspy et al., 2006), and early intervention services for first-episode psychosis. However, whatever the model of local service provision, their role in the provision of specialist clinical, advisory, training and gate-keeping functions will need to be clearly set out in the protocols governing the operation of the network. Typically, given expected demand for inpatient care, a network brings together a number of specialist perinatal teams (normally coterminous with a specialist mental health trust).

In a managed network, referral pathways for women requiring specialist care and sources of advice available to healthcare professionals without specialist training would be managed using protocols agreed within the network. In particular, a managed network should aim to provide:

- active working relationships between healthcare professionals working in different parts of the network
- shared care protocols
- shared educational and training programmes
- shared user groups or user group networks
- explicit pathways of care following a woman's journey through care.

Women identified by general medical services, such as maternity services or through their GPs, as having a mental health problem can then either be referred directly to the part of the network that can give them the most appropriate care, or healthcare professionals in general medical services can source appropriate information and advice from colleagues in other parts of the network to provide adequate care themselves. A crucial aspect of the network should be that it will provide for women with severe mental health problem, such as schizophrenia or bipolar disorder, prompt advice and, where appropriate, treatment from specialist perinatal mental health services, where necessary facilitating prompt access to specialist inpatient services.

4.4.2 Estimating need in the managed network model

The estimation of need in this model for the 2007 guideline review started with one of the building blocks of the network, the need for inpatient care. In section 4.2.2 the number of additional beds required was estimated at between 60 and 80. However, considerable variation of need and provision of existing services between the areas covered by the perinatal networks was estimated. The 2007 guideline recommended that each managed network should cover a population of between 25,000 and 50,000 live births, depending on local population morbidity. A key task outlined for the local networks was to determine need for all levels of care, including inpatient care, in light of the local epidemiology and current service provision and configuration.

4.5 RECOMMENDATIONS

4.5.1 Clinical recommendations

4.5.1.1 Clinical networks should be established for perinatal mental health services, managed by a coordinating board of healthcare professionals, commissioners, managers, and service users and carers. These networks should provide:

- a specialist multidisciplinary perinatal service in each locality, which provides direct services, consultation and advice to maternity services, other mental health services and community services; in areas of high morbidity these services may be provided by separate specialist perinatal teams
- access to specialist expert advice on the risks and benefits of psychotropic medication during pregnancy and breastfeeding
- clear referral and management protocols for services across all levels of the existing stepped-care frameworks for mental health problems, to ensure effective transfer of information and continuity of care
- pathways of care for service users, with defined roles and competencies for all professional groups involved. **[2007]**

4.5.1.2 Each managed perinatal mental health network should have designated specialist inpatient services and cover a population where there are between 25,000 and 50,000 live births a year, depending on the local psychiatric morbidity rates. **[2007]**

4.5.1.3 Specialist perinatal inpatient services should:

- provide facilities designed specifically for mother and babies (typically with 6–12 beds)
- be staffed by specialist perinatal mental health staff
- be staffed to provide appropriate care for babies
- have effective liaison with general medical and mental health services
- have available the full range of therapeutic services
- be closely integrated with community-based mental health services to ensure continuity of care and minimum length of stay. **[2007]**

4.5.1.4 Women who need inpatient care for a mental health problem within 12 months of childbirth should normally be admitted to a specialist mother and baby unit, unless there are specific reasons for not doing so. **[2007]**

4.5.1.5 Managers and senior healthcare professionals responsible for perinatal mental health services (including those working in maternity and primary care services) should ensure that:

- there are clearly specified care pathways so that all primary and secondary healthcare professionals involved in the care of women during pregnancy and the postnatal period know how to access assessment and treatment
- staff have supervision and training, covering mental health problems, assessment methods and referral routes, to allow them to follow the care pathways. **[2007]**

4.5.2 Research recommendations

4.5.2.1 Assessing managed perinatal networks

An evaluation of managed perinatal networks should be undertaken to compare the effectiveness of different network models in delivering care. It should cover the degree of integration of services, the establishment of common protocols, the impact on patients' access to specified services and the quality of care, and staff views on the delivery of care. **[2007]**

5 CASE IDENTIFICATION AND ASSESSMENT

5.1 INTRODUCTION

Pregnancy and the postnatal period are critical transitional periods for women. Culturally women expect the pregnancy and the birth of a new baby to be a positive and happy experience. However, for a significant number of women it can be a time of acute distress and illness, with a reluctance to admit how they are feeling because of the stigma that is associated with a failure to conform to the stereotype, and concerns that they might be regarded as being unfit to parent their baby (see Chapter 6).

Fathers may also experience mental health problems during their partner's pregnancy and the postnatal period, with a meta-estimate of prevalence in the region of 10%, rising to 25.6% in the 3 to 6 months after childbirth, and evidence of a moderate and positive correlation between maternal and paternal depression in the postnatal period (Paulson & Bazemore, 2010).

While the aetiology and course of mental health problems in pregnancy and the postnatal period are broadly the same as those that occur at other times, the different context in terms of the presence of a fetus and baby, have significant implications both in terms of identification and treatment.

Mental health problems in pregnancy and the postnatal period can have a significant impact on other family members including the woman's partner (Schumacher et al., 2008; Davey et al., 2006), but the most far-reaching consequences can occur in terms of the woman's relationship with her newborn baby, and the long-term development of the infant (see Chapter 7).

Although the early identification of women who are both at risk of or experiencing mental health problems in pregnancy and the postnatal period provides an important window of opportunity to reduce the impact of such problems on the long-term development of the child, many opportunities for such identification are missed, and around 50% of cases can go undetected (Ramsay, 1993).This may be due to the failure of many professionals to ask women about their mental health in the postnatal period.

This chapter reviews evidence for: (a) the effectiveness of methods to predict and identify mental health problems in women who are pregnant or in the first postnatal year; and (b) tools to assess the impact of such mental health problems on the mother-baby relationship.

5.2 CLINICAL REVIEW PROTOCOL (CASE IDENTIFICATION AND ASSESSMENT)

The review protocol summary, including the review question(s), information about the databases searched, and the eligibility criteria used for this section of the guideline, can be found in Table 9 (a complete list of review questions can be found in Appendix 8; further information about the search strategy can be found in Appendix 10; the full review protocols can be found in Appendix 9).

A systematic review of the literature (both primary studies and systematic reviews) was conducted to evaluate appropriate methods or instruments which are used to identify mental health problems in women who are pregnant or in the first postnatal year. For case identification (research question [RQ] 3.2), pooled diagnostic accuracy meta-analyses on the sensitivity and specificity of specific case identification instruments when compared with a DSM-IV or ICD-10 diagnosis were conducted (dependent on available data). In the absence of adequate data, it was agreed by the GDG that a narrative review of case identification instruments would be conducted and guided by a pre-defined list of consensus-based criteria (for example, the clinical utility of the instrument, administrative characteristics, and psychometric data evaluating its sensitivity and specificity).

For the purposes of the review of assessment, it was decided that a narrative synthesis of available evidence would be conducted, and in the absence of adequate data, a consensus-based approach to identify the key components of an effective assessment would be used.

Table 9: Clinical review protocol for the review of case identification instruments and assessment of mental health problems in women who are pregnant or the postnatal period

Component	Description
Review question(s)	**Case identification** What concerns and behaviours (as expressed by the woman, carer and family, or exhibited by the woman) should prompt any professional who comes into contact with a woman who is pregnant or in the first postnatal year to consider referral or further assessment for the presence of mental health problems? (RQ3.1) What are the most appropriate methods/ instruments for the identification of mental health problems in women who are pregnant or in the first postnatal year? (RQ3.2) **Assessment** For women who are pregnant or in the postnatal period, what are the key components of, and the most appropriate structure for a comprehensive diagnostic assessment (including diagnosis)? (RQ3.3)
Objectives	For case identification (RQ3.2) To identify brief screening instruments (< 12 items) which assess for mental health problems in women who are pregnant or in the postnatal period To assess the diagnostic accuracy of brief screening instruments.
Criteria for considering studies for the review	
Population	Women who are pregnant or in the postnatal period (from delivery to the end of the first year)
Intervention	For case identification (RQ3.2): brief screening instruments (< 12 items) for example, the Edinburgh Postnatal Depression Scale
Comparison	Gold standard: Diagnosis Statistical Manual (DSM-IV) or International Classification of Diseases (ICD-10)
Critical outcomes	**Sensitivity:** the proportion of true positives of all cases diagnosed with a mental health problem in the population **Specificity:** the proportion of true negatives of all cases not-diagnosed with a mental health problem in the population.
Important, but not critical outcomes	**Positive predictive value (PPV):** the proportion of patients with positive test results who are correctly diagnosed. **Negative predictive value (NPV):** the proportion of patients with negative test results who are correctly diagnosed. **Area under the curve (AUC):** constructed by plotting the true positive rate as a function of the false positive rate for each threshold.
Study design	Systematic reviews of RCTs Primary RCTs Cross sectional studies (including both cohort and case-control studies)
Include unpublished data?	No
Restriction by date?	No
Minimum sample size	No
Search strategy	**Databases searched**: General medical databases: Embase, Medline, PreMedline, PsycINFO

	Study design searched: All study designs **Date restrictions:** None, database inception to 07 April 2014
Searching other resources	Hand-reference searching of retrieved literature.

5.3 CASE IDENTIFICATION

5.3.1 Introduction

Women typically have frequent contact with a range of healthcare professionals during pregnancy, childbirth and the postnatal period, which presents an opportunity to identify those at risk of developing, or currently experiencing a mental health problem. However, identification rates are low; in the case of postnatal depression less than 50% of cases are identified by primary healthcare professionals in routine clinical practice (Hewitt et al., 2009). This section of the chapter assesses evidence for the effectiveness of instruments to identify mental health problems in pregnancy and the postnatal period.

Definition and aim of review

The review aims to identify and evaluate the diagnostic accuracy of brief case identification instruments for detecting mental health problems in women who are pregnant or the postnatal period.

For the purposes of this review, case identification instruments are defined as validated psychometric measures used to identify mental health problems in women in pregnancy or the postnatal period. This review was limited to instruments likely to be used in UK clinical practice that is, 'brief instruments', defined as those which are less than 12 items. 'Gold standard' diagnoses were defined as a DSM (American Psychiatric Association, 2013) or ICD (WHO, 1992) diagnosis; studies were sought that compared case identification using a brief instrument with a gold standard.

5.3.2 Methodological approach

The following criteria were considered when evaluating case identification instruments for inclusion in the review:

Quality of diagnostic test accuracy studies: the QUADAS-2 tool (a quality assessment tool for diagnostic accuracy studies; Whiting et al., 2011) was used to assess the quality of the evidence from diagnostic test accuracy studies. Each study was assessed for risk of bias (in terms of participant selection, the index test, and the reference standard) and for applicability (the extent to which the participant selection, index test and reference standard were applicable with regards to the review question). The GDG considered the quality assessment together with the

criteria listed below in making recommendations for case identification and assessment tools.

Primary aim of the instrument: the identification of mental health problems but not the formal diagnosis or the assessment of a particular disorder.

Clinical utility: the instrument should be feasible and implementable in routine clinical care. The instrument should contribute to the identification of further assessment needs and inform decisions about referral to other services.

Instrument characteristics and administrative properties: the case identification tool should have well-validated cut-offs in the population of interest. A case identification instrument should be brief, easy to administer and score, and be able to be interpreted without extensive and specialist training; it should also contain no more than 12 items. The instrument should be available in practice and free to use where possible.

Population: the population being assessed included any women who are pregnant or in the postnatal period up to 1 year. The review sought to assess screening tools used to detect mental health problems in pregnancy and the postnatal period across a variety of settings and in different languages of administration and did not limit instruments to those validated in a UK population.

Psychometric data: the instrument should have established reliability and validity (although these data will not be reviewed at this stage). It must have been validated against a gold standard diagnostic instrument such as DSM-IV or ICD-10 and it must have been reported in a paper that described its sensitivity and specificity.

Summary statistics used to evaluate identification instruments

Sensitivity and specificity

The terms 'sensitivity' and 'specificity' are used in relation to identification methods discussed in this chapter.

The **sensitivity** of an instrument refers to the proportion of those with the condition who test positive. An instrument that detects a low percentage of cases will not be very helpful in determining the numbers of patients who should receive a known effective treatment, as many individuals who should receive the treatment will not do so. This would lead to an under-estimation of the prevalence of the disorder, contribute to inadequate care and make for poor planning and costing of the need for treatment. As the sensitivity of an instrument increases, the number of false negative sit detects will decrease.

The **specificity** of an instrument refers to the proportion of those who do not have the condition and test negative. This is important so that healthy people are not

offered treatments they do not need. As the specificity of an instrument increases, the number of false positives will decrease.

To illustrate this, from a population in which the point prevalence rate of depression is 10% (that is, 10% of the population has depression at any one time), 1,000people are given a test which has 90% sensitivity and 85% specificity. It is known that100 people in this population have depression, but the test detects only 90 (true positives), leaving 10 undetected (false negatives). It is also known that 900 people do not have depression, and the test correctly identifies 765 of these (true negatives), but classifies 135 incorrectly as having depression (false positives). The positive predictive value of the test (the number correctly identified as having depression as a proportion of positive tests) is 40% (90/90 + 135), and the negative predictive value (the number correctly identified as not having depression as a proportion of negative tests) is 98% (765/765 +10). Therefore, in this example, a positive test result is correct in only 40% of cases, while a negative result can be relied upon in 98% of cases.

The example above illustrates some of the main differences between positive predictive values and negative predictive values in comparison with sensitivity and specificity. For both positive and negative predictive values, prevalence explicitly forms part of their calculation (see Altman & Bland, 1994a). When the prevalence of a disorder is low in a population this is generally associated with a higher negative predictive value and a lower positive predictive value. Therefore although these statistics are concerned with issues probably more directly applicable to clinical practice (for example, the probability that a person with a positive test result actually has depression), they are largely dependent on the characteristics of the population sampled and cannot be universally applied (Altman & Bland, 1994a).

On the other hand, sensitivity and specificity do not necessarily depend on prevalence of depression (Altman & Bland, 1994b). For example, sensitivity is concerned with the performance of an identification test conditional on a person having depression. Therefore the higher false positives often associated with samples of low prevalence will not affect such estimates. The advantage of this approach is that sensitivity and specificity can be applied across populations (Altman & Bland, 1994b). However, the main disadvantage is that clinicians tend to find such estimates more difficult to interpret.

When evaluating diagnostic accuracy, sensitivity and specificity were used as the most suitable summary statistics due to the fact that the studies included were from a range of populations, included both cohort and case-control designs, and populations where mother were 'at risk' of mental health problems, therefore resulting in variations in prevalence.

When describing the sensitivity and specificity of the different instruments, the GDG defined values above 0.9 as 'excellent', 0.8 to 0.9 as 'good', 0.5 to 0.7 as moderate', 0.3 to 0.5 as 'low', and less than 0.3 as 'poor'.

Receiver operating characteristic curves

The qualities of a particular tool are summarised in a receiver operating characteristic (ROC) curve, which plots sensitivity (expressed as a per cent) against (100-specificity).

A test with perfect discrimination would have a ROC curve that passed through the top left-hand corner; that is, it would have 100% specificity and pick up all true positives with no false positives. While this is never achieved in practice, the area under the curve (AUC) measures how close the tool gets to the theoretical ideal. A perfect test would have an AUC of 1, and a test with AUC above 0.5 is better than chance. As discussed above, because these measures are based on sensitivity and 100-specificity, theoretically these estimates are not affected by prevalence.

5.3.3 Studies considered[7]

Case identification instruments included in the review

There were four instruments which met the inclusion criteria for case identification which are included in the review: the Edinburgh Postnatal Depression Scale (EPDS; Cox et al., 1987); the Patient Health Questionnaire (PHQ; Spitzer et al., 1999); the 'Whooley questions' (Whooley et al., 1997); and the Kessler-10 (Kessler et al., 2002). The mental health problems evaluated by these instruments were depression and, or, anxiety. Study characteristics for case identification tools included in the review can be found in Table 10. To maximise the available data, the most consistently reported and recommended cut-off points for each of the scales were extracted

Results of the search

To be included in the review, a study must have reported the sensitivity and specificity of the instrument relative to a diagnostic interview for the relevant cut-off points, or sufficient data were available for these parameters to be calculated. Studies that did not clearly state the comparator to be diagnosis by DSM or ICD, used a scale with greater than 12 items, or did not provide sufficient data to be included in the review were excluded. To be included in the meta-analyses the studies must have reported enough information to calculate the true positives, true negatives, false positives and false negatives.

The literature search for observational studies yielded 9,897 articles overall. Scanning titles or abstracts identified 122 potentially relevant studies that evaluated the recognition and case identification of mental health problems in women who are pregnant or in the postnatal period.

[7]Here and elsewhere in the guideline, each study considered for review is referred to by a study ID in capital letters (primary author and date of study publication, except where a study is in press or only submitted for publication, then a date is not used).

After further inspection of the full citations, 51 studies did not meet one or more eligibility criteria. The most common reasons for exclusion were: studies reported on instruments with more than 12 items, there was no suitable gold standard tool, studies did not have relevant outcomes (for example, did not provide sensitivity and specificity data), the studies were not in English or the population was not relevant.

A further study (KADIR2005 [Kadir et al., 2005]) was identified from hand-searches of relevant articles yielding a total of 72 studies overall. In addition, a systematic review of validation studies for the EPDS was identified, GIBSON2009 (Gibson et al., 2009) which was used as a source of data from two studies where there was no access to the full papers (ASCASO2003 [AscasoTerren et al., 2003], JADRESIC1995 [Jadresic et al., 1995]). Further information about both included and excluded studies can be found in Appendix 18. A summary of the methodological quality of the included studies can be found in Figure 3, and the full methodological checklists can be found in Appendix 17.

As a result of this, a total of 72 published studies met the eligibility criteria for this review, however only 60 studies provided sufficient data to be included in the statistics analysis: ADEWUYA2005 (Adewuya et al., 2005), ADEWUYA2006 (Adewuya et al., 2006), AGOUB2005 (Agoub et al., 2005), ALVARADO-ESQUIVEl2006 (Alvarado-Esquivel et al., 2006), ASCASO2003, AYDIN2004 (Aydin et al., 2004), BAGGALEY2007 (Baggaley et al., 2007), BARNETT1999 (Barnett et., 1999), BECK2001 (Beck et al., 2001), BENVENUTI1999 (Benvenuti et al., 1999), BERGINK2011 (Bergink et al., 2011), BERLE2003 (Berle et al., 2003), BOYCE1993 (Boyce et al., 1993), BUNEVICIUS2009 (Bunevicius et al., 2009), CARPINIELLO1997 (Carpiniello et al., 1997), CHAUDRON2010 (Chaudron et al., 2010), CHIBANDA2010 (Chibanda et al., 2010), CLARKE2008 (Clarke et al., 2008), COX1987 (Cox et al., 1987), EBERHARD-GRAN2001 (Eberhard-Gran et al., 2001), EKEROMA2012 (Ekeroma et al., 2012), FELICE2006 (Felice et al., 2006), FERNANDES2011 (Fernandes et al., 2011), FLYNN2011 (Flynn et al., 2011), GARCIA-ESTEVE2003 (Garcia-Esteve et al., 2003), GAUSIA2007 (Gausia et al., 2007), GHUBASH1997 (Ghubash et al., 1997), GJERDINCJEN2009 (Gjerdincjen et al., 2009), GUEDENEY1998 (Guedeney et al., 1998), HARRIS1989 (Harris et al., 1998), JADRESIC1995, KADIR2005, LAU2010 (Lau et al., 2010), LEE1998 (Lee et al., 1998), LEONARDOU2009 (Leonardou et al., 2009), LEVERTON2000 (Leverton et al., 2000), MAHMUD2003 (Mahmud et al., 2003), MANN2012 (Mann et al., 2012), MATTHEY2008 (Matthey et al., 2008), MAZHARI2007 (Mazhari et al., 2007), MILGROM2005A (Milgrom et al., 2005a), MURRAY1990B (Murray et al., 1990b), MUZIK2000 (Muzik et al., 2000), PHILLIPS2009 (Phillips et al., 2009), PITANUPONG2007 (Pitanupong et al., 2007), REGMI2002 (Regmi et al., 2002), RUBERTSSON2011 (Rubertsson et al., 2011), SANTOS2007 (Santos et al., 2007), SIDEBOTTOM2012 (Sidebottom et al., 2012), SMITH2010 (Smith et al., 2010), SPIES2009 (Spies et al., 2009), TANDON2012 (Tandon et al., 2012), TENG2005 (Teng et al., 2005), THIAGAYSON2013 (Thiagayson et al., 2013), TOREKI2013 (Toreki et al., 2013), TRAN2011 (Tran et al., 2011), UWAKWE2003 (Uwakwe et al., 2003), WERRETT2006 (Werrett et al., 2006), WICKBERG1996 (Wickberg et al., 1996), YOSHIDA2001 (Yoshida et al., 2001).

Twelve studies met the inclusion criteria but were not included in the meta-analysis because the data could not be extracted or the population was not appropriate for the cut-off points used: AREIAS1996 (Areias et al., 1996), HANLON2008 (Hanlon et al., 2008), HANUSA2008 (Hanusa et al., 2008), JARDRI2006 (Jardri et al., 2006), JI2006 (Ji et al., 2011) LAWRIE1998A (Lawrie et al., 1998a), LOGSDON2010 (Logsdon et al., 2010) MURRAY1990A (Murray et al., 1990a), ROWEL2008 (Rowel et al., 2008), STEWART2013 (Stewart et al., 2013), VENKATESH2013 (Venkatesh et al., 2013) ZELKOWITZ1995 (Zelkowitz et al., 1995).

Of the eligible studies, here were 54 which were included in the meta-analysis for the EPDS (Table 11), four included the meta-analysis for the PHQ (Table 12), two included in the meta-analysis for the Whooley questions (Table 13), and three studies for the Kessler-10 (Table 14). Two of these studies (BARNETT1999, EKEROMA2012) reported data on more than one population.

Table 10: Characteristics of case identification instruments included in the review

Instrument	Mental health problem evaluated	Population	Number of items (scale)	Completed by Version	Time to administer and score/training required/cost and copyright issues
EPDS	Depression (and anxiety)	Women of child bearing age	10 items (0–30)	Self-report Pen and paper format	Administration time: 10 minutes Scoring time: 5 minutes Training Support: none described, but none seems to be needed Freely available
PHQ	Depression	All adults (mainly used in primary care settings)	9 items (0-27) 8 items (0-24) 2 items (0-6)	Self-report Pen and paper format	Administration time: depending on tool, 3–10 minutes Scoring Time: 5 minutes Training support: experienced clinician Freely available
Kessler-10	Depression and anxiety	All adults	10 items (0-50)	Self-report Pen and Paper	Administration time: 10 minutes Scoring time: 5 minutes Training Support: none described Freely available
Whooley questions	Depression (and anxiety)	All adults	2 items (plus help question)	Self-report verbal, telephone	Administration Time: < minute Scoring Time: < minute

			Yes/No response		Training Support: none described Freely available

Table 11: Study information table for studies included in the review for the EPDS

Study ID K=54 (57 populations)	N	Study design	Country	Language	Mean age (years)	Timing	Identified risk factors	Diagnosis	Index cut-off
ADEWUYA2005	876	Cohort	Nigeria	English or Yoruba	29	Postnatal	No	Major depression; Mixed depression	9/10 12/13
ADEWUYA2006	182	Case-control	Nigeria	Nigeria	25	Pregnancy	No	Major depression; Mixed depression	9/10 12/13
AGOUB2005	144	Cohort	Nigeria	Arabic	30	Postnatal	No	Mixed depression	9/10 12/13
ALVARADO-ESQUIVEl2006	100	Cohort	Mexico	Mexican	24	Postnatal	Yes	Mixed depression	9/10 12/13
ASCASO2003	334	Cohort	Spain	Spain	25	Pregnancy and postnatal	No	Mixed depression	9/10 12/13
AYDIN2004	341	Cohort	Turkey	Turkish	27	Postnatal	No	Mixed depression	9/10 12/13
BARNETT1999(A)	98	Cohort	Australia	Arabic	NR	Postnatal	No	Major depression	9/10 12/13
BARNETT1999(AC)	105	Cohort	Australia	Anglo-Celtic	NR	Postnatal	No	Major depression	9/10 12/13
BARNETT1999(V)	113	Cohort	Australia	Vietnamese	NR	Postnatal	No	Major depression	9/10 12/13
BECK2001	150	Cohort	US	English	31	Postnatal	No	Mixed depression	9/10 12/13
BENVENUTI1999	32	Cohort	Italy	Italian	32	Postnatal	No	Major depression; Mixed depression	9/10 12/13

BERGINK2011	854	Cohort	Netherlands	Dutch	30	Pregnancy	No	Major depression	9/10 12/13
BERLE2003	100	Case-control	Norway	Norwegian	30	Postnatal	Yes	Major depression; Mixed depression	9/10 12/13
BOYCE1993	103	Case-control	Australia	English	28	Postnatal	No	Major depression	9/10 12/13
BUNEVICIUS2009	230	Cohort	Lithuania	Lithuanian	29	Pregnancy	No	Mixed depression	12/13
CARPINIELLO1997	61	Cohort	Italy	Italian	32	Postnatal	No	Mixed depression	9/10 12/13
CHAUDRON2010	61	Cohort	US	English	32	Postnatal	Yes	Mixed depression	9/10 12/13
CHIBANDA2010	210	Cohort	Zimbabwe	Shona (local language)	25	Postnatal	No	Major depression	9/10 12/13
CLARKE2008	103	Cohort	Canada	English	24	Postnatal	No	Mixed depression	12/13
COX1987	96	Case-control	UK	English	24	Postnatal	No	Mixed depression	12/13
EBERHARD-GRAN2001	56	Case-control	Norway	Norwegian	30	Postnatal	No	Major depression	9/10
EKEROMA2012(T)	85	Cohort	New Zealand	Tongan	29	Postnatal	No	Major depression	9/10 12/13
EKEROMA2012(S)	85	Cohort	New Zealand	Samoan	30	Postnatal	No	Major depression	9/10 12/13
FELICE2006	233	Cohort	Malta	Maltese	27	Pregnancy and postnatal	No	Mixed depression	9/10 12/13 14/15
FERNANDES2011[1]	194	Cohort	India	Indian	22	Pregnancy	No	Mixed depression	9/10 12/13 14/15
FLYNN2011[2]	185	Cohort	US	English	30	Pregnancy and postnatal	No	Major depression	12/13
GARCIA-ESTEVE2003	334	Cohort	Spain	Spanish	30	Pregnancy and postnatal	No	Major depression; Mixed depression	9/10 12/13
GAUSIA2007	126	Cohort	Bangladesh	Bengali	26	Postnatal	No	Mixed depression	9/10 12/13

GHUBASH1997	95	Cohort	United Arab Emirates	Arabic	29	Postnatal	No	Mixed depression	9/10 12/13
GUEDENEY1998	87	Case-control	France	French	30	Postnatal	Yes	Mixed depression	9/10 12/13
HARRIS1989	126	Cohort	UK	English		Postnatal	No	Major depression	12/13
JADRESIC1995	108	Cohort	Chile	Spanish	28	Postnatal	No	Mixed depression	9/10 12/13
KADIR2005	52	Cohort	Malaysia	Malay	NR	Postnatal	No	Major depression; Mixed depression	9/10 12/13
LAU2010	342	Cohort	China	Chinese	NR	Postnatal	No	Mixed depression	9/10 12/13
LEE1998	145	Cohort	Hong Kong	Chinese	29	Postnatal	No	Mixed depression	9/10 12/13
LEONARDOU2009	81	Cohort	Greece	Greek	32	Postnatal	No	Mixed depression	9/10 12/13
LEVERTON2000	199	Cohort	UK	English	NR	Postnatal	No	Mixed depression	9/10 12/13
MAHMUD2003	64	Cohort	Malaysia	Malay	29	Postnatal	No	Mixed depression	9/10 12/13
MATTHEY2008	238	Cohort	Australia	English	27	Postnatal	No	Anxiety disorder	3/4 4/5 5/6
MAZHARI2007	200	Case-control	Iran	Farsi	26	Postnatal	Yes	Major depression; Mixed depression	9/10 12/13
MILGROM2005A	344	Cohort	Australia	English	30	Postnatal	Yes	Mixed depression	12/13
MURRAY1990B	100	Cohort	UK	English	NR	Pregnancy	No	Major depression; Mixed depression	12/13 14/15
MUZIK2000	50	Cohort	Austria	German	28	Postnatal	No	Major depression	9/10 12/13
PHILLIPS2009	166	Cohort	Australia	English	32	Postnatal	No	Major depression; Anxiety disorders	3/4 4/5 5/6 12/13
PITANUPONG2007	615	Cohort	Thailand	Thai	28	Postnatal	No	Mixed depression	9/10 12/13

REGMI2002	140	Case-control	Nepal	Nepali	NR	Postnatal	No	Major depression	12/13
RUBERTSSON2011	121	Cohort	Sweden	Swedish	30	Pregnancy	No	Major depression	12/13
SANTOS2007	378	Case-control	Brazil	Portuguese	NR	Postnatal	Yes	Mixed depression	9/10 12/13
TANDON2012	92	Cohort	US	English	24	Postnatal	Yes	Major depression; Mixed depression	9/10 12/13
TENG2005	203	Cohort	Taiwan	Taiwanese	29	Postnatal	No	Mixed depression	12/13
THIAGAYSON2013	200	Cohort	Singapore	NR	31	Pregnancy and postnatal	No	Major depression; Mixed depression; Anxiety disorders	8/9 9/10 12/13
TOREKI2013	219	Cohort	Hungary	Hungarian	30	Pregnancy	No	Major depression; Mixed depression	9/10 12/13 14/15
TRAN2011	364	Cohort	Vietnam	Vietnamese	NR	Pregnancy and postnatal	No	Common mental health disorder	3/4 4/5 5/6
UWAKWE2003	225	Cohort	Nigeria	Igbo	29	Postnatal	No	Mixed depression	9/10 12/13
WERRETT2006	23	Cohort	Asian	English and Punjabi	29	Postnatal	No	Mixed depression	9/10 12/13
WICKBERG1996	41	Case-control	Sweden	Swedish	28	Postnatal	No	Major depression	12/13
YOSHIDA2001	98	Cohort	UK/Japan	Japanese	NR	Postnatal	No	Mixed depression	9/10 12/13

Note. Abbreviations: NR=not reported.
[1] FERNANDES2011 reports data for both the EPDS and Kessler-10.
[2] FLYNN2011 reports data for both the EPDS and PHQ.

Table 12: Study information table for studies included in the review for the PHQ

Study ID K=4	N	Study design	Country	Language	Mean age (years)	Timing	Identified risk factors	Diagnosis	Index cut-off
FLYNN2011[1]	185	Cohort	US	English	30	Pregnancy and postnatal	No	Major depression	9/10
GJERDINCJEN2009[2]	506	Cohort	US	English	29	Postnatal	N/A	Major depression	9/10
SIDEBOTTOM2012	745	Cohort	US	English	23	Pregnancy	N/A	Major depression; Mixed depression	9/10
SMITH2010 (PHQ-9 and -2)	218	Cohort	US	English	29	Pregnancy	N/A	Major depression	3/4 9/10
Note. Abbreviations: NR=not reported. [1] FLYNN2011 reports data for both the EPDS and PHQ. [2] GJERDINCJEN2009 reports data for both the PHQ and Whooley questions.									

Table 13: Study information table for studies included in the review of the Whooley questions

Study ID K=2	N	Study design	Country	Language	Mean age (years)	Timing	Identified risk factors	Diagnosis	Index cut-off
GJERDINCJEN2009[1]	506	Cohort	US	English	29	Postnatal	No	Major depression	N/A
MANN2012	152	Cohort	UK	English	27	Pregnancy and postnatal	No	Major Depression	N/A
Note. Abbreviations: NR=not reported									

[1] GJERDINCJEN2009 reports data for both the PHQ and Whooley questions

Table 14: Study information table for studies included in the review of the Kessler-10

Study ID K=3	N	Study design	Country	Language	Mean age (years)	Timing	Identified risk factors	Diagnosis	Index cut-off
BAGGALEY2007	61	cohort	Burkina Faso	West African French and local languages	26	Postnatal	Yes	Mixed depression	5/6
FERNANDES2011[1]	194	cohort	India	Indian	22	Postnatal	No	Mixed depression	5/6
SPIES2009	129	cohort	South Africa	Afrikaans	NR	Pregnancy	No	Anxiety disorders	5/6

Note. Abbreviations: NR=not reported
[1] FERNANDES2011 reports data for both the EPDS and Kessler-10

Figure 3: Methodological quality of studies included in the review

Study ID	Index test	Risk of bias				Applicability concerns		
		Patient selection	Index test	Reference standard	Flow and timing	Patient selection	Index test	Reference standard
ADEWUYA2005	EPDS	+	?	-	?	-	+	-
ADEWUYA2006	EPDS	-	-	+	-	-	-	+
AGOUB2005	EPDS	+	?	?	?	+	+	+
ALVARADO-ESQUIVEl2006	EPDS	+	+	+	+	-	+	+
AYDIN2004	EPDS	+	+	+	+	+	+	+
BAGGALEY2007	Kessler-10	+	+	+	+	+	+	?
BARNETT1999(A)	EPDS							
BARNETT1999(AC)								
BARNETT1999(V)		+	?	?	+	+	-	+
BECK2001	EPDS	+	+	+	?	+	+	+
BENVENUTI1999	EPDS	+	?	+	?	+	+	+
BERGINK2011	EPDS	+	?	+	-	+	+	+
BERLE2003	EPDS	-	+	+	-	+	+	+
BOYCE1993	EPDS	-	+	?	?	-	+	+
BUNEVICIUS2009	EPDS	+	+	+	?	+	+	+
CARPINIELLO1997	EPDS	+	+	?	+	+	+	+
CHAUDRON2010	EPDS	+	?	+	-	+	+	+
CHIBANDA2010	EPDS	+	+	+	+	+	+	+
CLARKE2008	EPDS	+	?	?	?	+	+	+
COX1987	EPDS	-	-	+	?	+	+	+
EBERHARD-GRAN2001	EPDS	-	?	+	-	+	+	+
EKEROMA2012(T)	EPDS							
EKEROMA2012(S)		+	+	+	-	+	+	+
FELICE2006	EPDS	+	+	+	+	+	+	+
FERNANDES2011	EPDS	+	?	?	+	-	-	+
FLYNN2011	EPDS	+	+	?	-	+	+	-

Study ID	Index test	Risk of bias				Applicability concerns		
		Patient selection	Index test	Reference standard	Flow and timing	Patient selection	Index test	Reference standard
	PHQ							
GARCIA-ESTEVE2003	EPDS	-	+	+	-	+	+	+
GAUSIA2007	EPDS	+	+	+	+	+	?	+
GHUBASH1997	EPDS	+	?	?	?	+	+	+
GJERDINCJEN2009	PHQ, Whooley	+	?	?	?	+	+	+
GUEDENEY1998	EPDS	-	+	+	-	+	+	+
HARRIS1989	EPDS	+	?	+	+	+	+	+
KADIR2005	EPDS	+	?	?	?	+	+	+
LAU2010	EPDS	+	?	+	+	+	+	+
LEE1998	EPDS	+	+	+	-	+	+	+
LEONARDOU2009	EPDS	+	?	+	?	+	+	+
LEVERTON2000	EPDS	-	+	+	-	+	+	+
MAHMUD2003	EPDS	+	+	+	+	+	+	+
MANN2012	Whooley	+	+	+	-	+	+	+
MATTHEY2008	EPDS	+	?	+	?	+	?	+
MAZHARI2007	EPDS	-	+	+	-	+	-	+
MILGROM2005A	EPDS	-	+	?	-	+	+	+
MURRAY1990B	EPDS	+	?	+	?	+	+	+
MUZIK2000	EPDS	-	?	?	-	+	?	+
PHILLIPS2009	EPDS	+	?	+	-	+	+	+
PITANUPONG2007	EPDS	+	?	+	-	+	+	+
REGMI2002	EPDS	-	?	?	-	+	?	?
RUBERTSSON2011	EPDS	+	?	?	-	+	+	+
SANTOS2007	EPDS	-	?	+	-	+	+	+
SIDEBOTTOM2012	PHQ	+	+	?	-	+	+	?
SMITH2010	PHQ	-	+	?	-	+	+	+
SPIES2009	Kessler-10	+	?	?	?	+	-	+
TANDON2012	EPDS	+	+	-	+	+	-	+

Study ID	Index test	Risk of bias				Applicability concerns		
		Patient selection	Index test	Reference standard	Flow and timing	Patient selection	Index test	Reference standard
TENG2005	EPDS	+	?	+	-	+	+	+
THIAGAYSON2013	EPDS	+	?	+	+	?	+	+
TOREKI2013	EPDS	+	+	+	+	+	+	+
TRAN2011	EPDS	+	+	+	+	+	+	+
UWAKWE2003	EPDS	+	+	?	-	+	-	+
WERRETT2006	EPDS	+	+	+	+	+	+	+
WICKBERG1996	EPDS	-	+	+	-	+	+	+
YOSHIDA2001	EPDS	+	+	?	?	+	?	+
Note. Risk of bias assessment was not possible for ASCASO2003 and JADRESIC1995 because full text was not available. Results were taken from GIBSON2009.								

5.3.4 Clinical evidence for case identification instruments for mental health problems in women who are pregnant or in the postnatal period

Review Manager 5 was used to summarise diagnostic accuracy data from each study using forest plots and summary ROC plots. Where more than two studies reported appropriate data, a bivariate diagnostic accuracy meta-analysis was conducted using Metadisc (Zamora et al., 2006) publically available at www.hrc.es/investigacion/metadisc_en.htm, in order to obtain pooled estimates of sensitivity, specificity using a random effects model. Pooled estimates were provided with their respective confidence intervals. Forest plots and ROC curves generated by Review Manager were also inspected in order to obtain a general overview of the accuracy estimates from each study. Metadisc allowed an exploration of heterogeneity using a statistical test for I^2. Heterogeneity was also explored by visual inspection of forest plot confidence intervals of accuracy estimates.

Heterogeneity is usually much greater in meta-analyses of diagnostic accuracy studies compared with RCTs (Cochrane Collaboration, 2012; Gilbody et al., 2007). Therefore, a higher threshold for acceptable heterogeneity in such meta-analyses is required. However where substantial heterogeneity existed, or when pooling studies resulted in I^2 >90%, additional subgroup analyses were conducted for possible factors that might influence accuracy estimates. The reasons for such heterogeneity were explored by relating study level covariates: country (developed or developing); study design (cohort or case-control); and population (risk factors for a mental health problem or no risk factors).

Evaluating identification instruments for depression

When evaluating instruments, separate analyses were conducted depending on:

- The type of depression diagnosis that the gold standard diagnostic interview was used to classify; some studies used a combination category of both 'minor and major depression' (hereafter referred to 'mixed depression') in the definition of depression whilst others used a stricter definition of major depression only.
- The timing at which the instrument was administered; in pregnancy or in the postnatal period.
- The cut-off point chosen to indicate a positive test; threshold effects can create a potential source of heterogeneity, therefore studies were pooled which used the most consistently reported and recommended cut-off points.

Edinburgh Postnatal Depression Scale

The EPDS (Cox et al., 1987) is a ten-item self-report questionnaire developed to assist professionals to identify depression in the postnatal period. It was developed in an attempt to address the problem of the pregnancy or postnatal status *per se* affecting

experiences typically taken as indicators of depression, such as disturbances in appetite, and is routinely administered to women at 6 to 8 weeks after childbirth by their health visitor. Based on existing literature, the most consistently reported and recommended cut-off points for the EPDS are 9/10 and 12/13 (Gibson et al., 2009) for detecting 'possible depression' and 'probable depression' respectively (Cox et al., 1987). In pregnancy a higher cut-off of 14/15 has been suggested (Murray & Cox, 1990). Studies were included if they provided extractable data for these cut-off points.

Of the eligible studies there were 66 which assessed the EPDS. Of these, 53 studies across 56 different populations included sufficient data to be included in the statistical meta-analysis. There were 13 studies which reported sensitivity and specificity but did not report enough information to calculate true positives, false positives, true negatives and false negatives, and two studies which used a population that was not appropriate at the relevant cut-off points and therefore not included in the meta-analyses.

Studies were undertaken in 34 different countries, and 14 of these studies used an English-language (rather than a translated) version of the EPDS. There were 26 studies which included assessment for both minor and major depression in the definition of depression, 17 studies for major depression only and 10 studies provided data for both definitions of depression.

Meta-analyses were conducted separately for the different cut-off points and definitions of depression. This yielded a two-by-two table for pooled sensitivity and specificity estimates for postnatal depression and two-by-three table for pooled sensitivity and specificity estimates of depression in pregnancy.

EPDS – Detection of depression in pregnancy

The EPDS has been less well validated in screening for depression during pregnancy compared to the postnatal period, and the cut-off values have been found to differ from the postnatal ones. The original UK study validating the EPDS in pregnancy (Murray & Cox, 1990) found that at the 12/13 cut-off rate, the EPDS had a sensitivity of 100% for major depression and a specificity of 87%, however specificity was improved to 96% at the cut-off 14/15, suggesting a higher cut-off was required to use the EPDS to detect depression in pregnancy. However it was noted that subsequent studies suggest a lower cut-off should be used (Bergink et al., 2011). Pooled sensitivity and specificity estimates were therefore calculated for the cut-off 14/15 in addition to 9/10 and 12/13.

There were 10 eligible studies validating the EPDS for detecting depression in pregnancy across the three cut-off points; five studies reported sensitivity and specificity of detecting mixed depression and nine studies for major depression only. Of the eligible studies there was one which used a case-control design and two studies administered to 'at risk' women. Two studies were from developing countries, and two used English language versions. Table 15 summarises the results

of the meta-analyses in terms of pooled sensitivity and specificity estimates and the range of test data across the included studies at the different cut-offs for detecting mixed depression and major depression only. See forest plots and summary ROC curves in Appendix 19 for individual data by study, and the full methodological checklists in Appendix 17. There was relatively high heterogeneity across all the analyses. This existed after conducting subgroup analyses by study-design, population and country.

Table 15: Evidence summary table for the EPDS administered in pregnancy

Diagnosis	Cut off	No of Participants (studies)	Sensitivity		Specificity	
			Pooled Sensitivity (95% CI)	Range of test data	Pooled Specificity (95% CI)	Range of test data
Mixed (major and minor) depression	9/10	728 (4)	0.74 (0.65-0.82)	0.5-0.75	0.86 (0.83-0.89)	0.77-0.97
	12/13	722 (4)	0.61 (0.5-0.72)	0.18-0.86	0.94 (0.92-0.96)	0.90-1.0
	14/15	542 (3)	0.47 (0.35-0.60)	0.14-0.66	0.98 (0.97-0.99)	0.97-1.0
Major depression	9/10	1,258 (3)	0.88 (0.89-0.94)	0.43-1.00	0.88 (0.86-0.90)	0.48-0.93
	12/13	1,219 (8)	0.83 (0.76-0.88)	0.29-1.00	0.90 (0.88-0.92)	0.73-0.99
	14/15	599 (4)	0.72 (0.58-0.84)	0.29-1.00	0.97 (0.95-0.98)	0.93-0.99

EPDS – detection of depression in the postnatal period

Of the eligible studies, there were 43 which validated the EPDS in the postnatal period; 28 were conducted in developed countries of which 12 used an English language version.

Table 16 and Figure 4 summarise the results of the meta-analyses in terms of pooled sensitivity and specificity estimates and the range of test data across the included studies at the cut-off scores 9/10 and 12/13 for detecting mixed depression and major depression only. See forest plots in Appendix 19 for individual data by study and the full methodological checklists in Appendix 17

There were 29 studies validating the EPDS in the postnatal period which used the cut-off point 9/10 to detect mixed depression. Visual inspection of the summary ROC curve (Figure 4) demonstrated a wide variation of data from individual studies. Pooled estimates were good for both sensitivity and specificity although there was very high heterogeneity for pooled specificity estimates (I^2=96.2%) which existed after conducting subgroup analyses by study-design, population and country. However, visual inspection of the summary ROC curves, subgrouped by

women with and without risk factors for depression (Figure 5), suggested better diagnostic accuracy for studies conducted in the population with no risk factors (and could be one potential source of heterogeneity).

There were 27 studies validating the EPDS using the cut-off point 12/13 for detecting mixed depression. The EPDS was found to have a moderate pooled sensitivity although there was high heterogeneity. The pooled specificity was excellent but heterogeneity very high (I^2=94.4%) and existed after conducting subgroup analyses by study-design, population and country type. However, visual inspection of the summary ROC curve (Figure 5) demonstrated a similar pattern of better diagnostic accuracy for populations not at risk of depression as with the lower cut-off.

There were 13 studies using the cut-off point 9/10 for detecting major depression in the postnatal period. This was after removing one study from the analysis (LOGSDON2010) as an adolescent population was used where the cut-off point was not deemed appropriate. The EPDS was found to have excellent sensitivity with moderate heterogeneity and good pooled specificity although relatively high heterogeneity (I^2=85.1%). Using the cut-off point 12/13 for detecting major depression there were 23 studies. The EPDS had good pooled sensitivity with relatively high heterogeneity and excellent pooled specificity although high heterogeneity (I^2=90.3%).

Table 16: Evidence summary table for the EPDS administered in the postnatal period

Diagnosis	Cut-off	No. of Participants (studies)	Sensitivity		Specificity	
			Pooled sensitivity (95% CI)	Range of test data	Pooled specificity (95% CI)	Range of test data
Mixed depression	9/10	5,463 (29)	0.83 (0.81-0.86)	0.59- 1.0	0.85 (0.84-0.86)	0.47-0.99
	12/13	5,209 (29)	0.68 (0.66-0.71)	0.34- 0.96	0.92 (0.92-0.93)	0.71- 1.0
Major depression	9/10	2,277 (13)	0.95 (0.92-0.97)	0.71- 1.0	0.82 (0.80-0.84)	0.62- 0.89
	12/13	4,355 (22)	0.80 (0.77-0.83)	0.55-1.0	0.93 (0.92-0.94)	0.52-0 .99

Figure 4: Summary of ROC curve for the EPDS administered in the postnatal period at different cut-off points and diagnoses

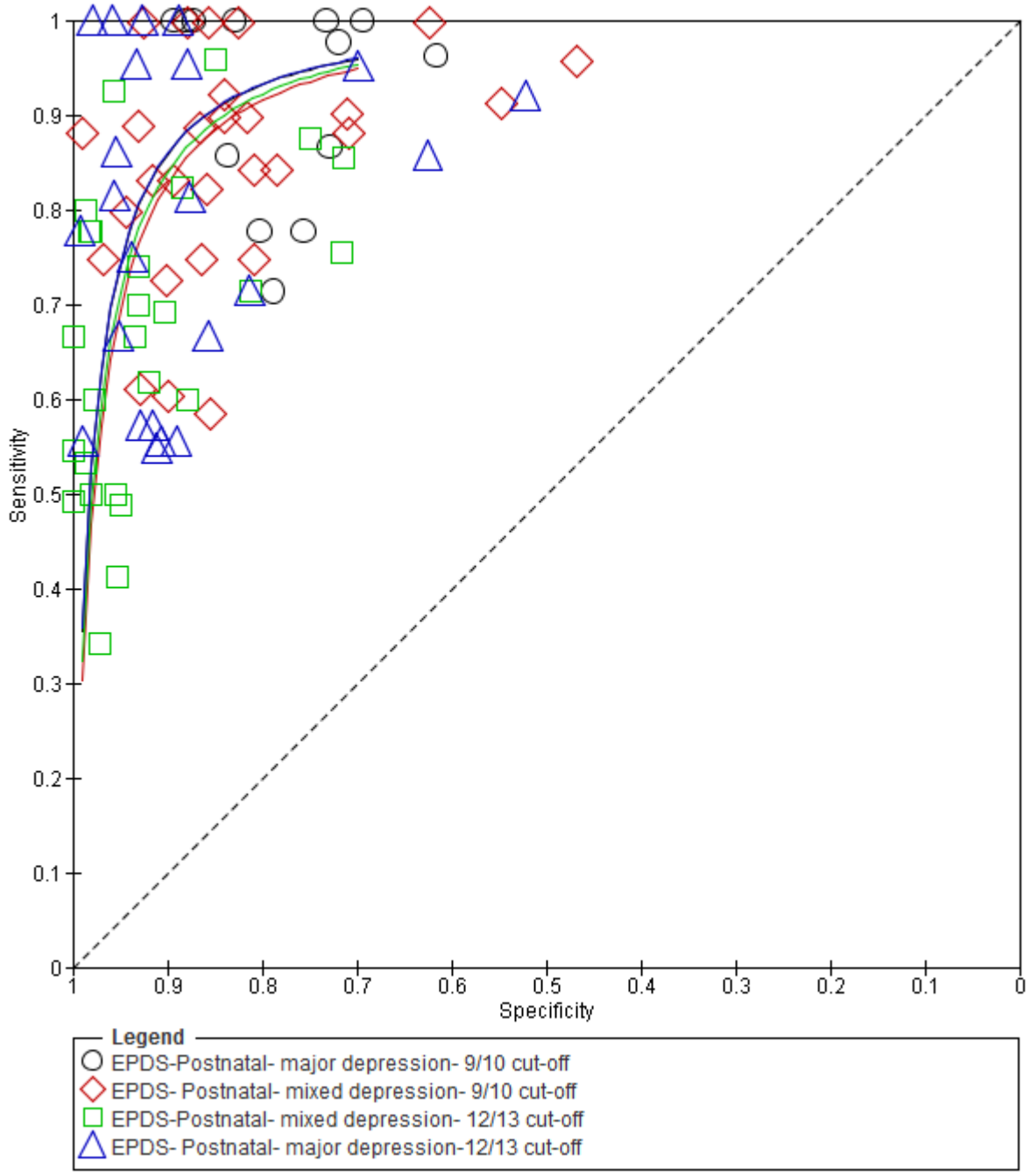
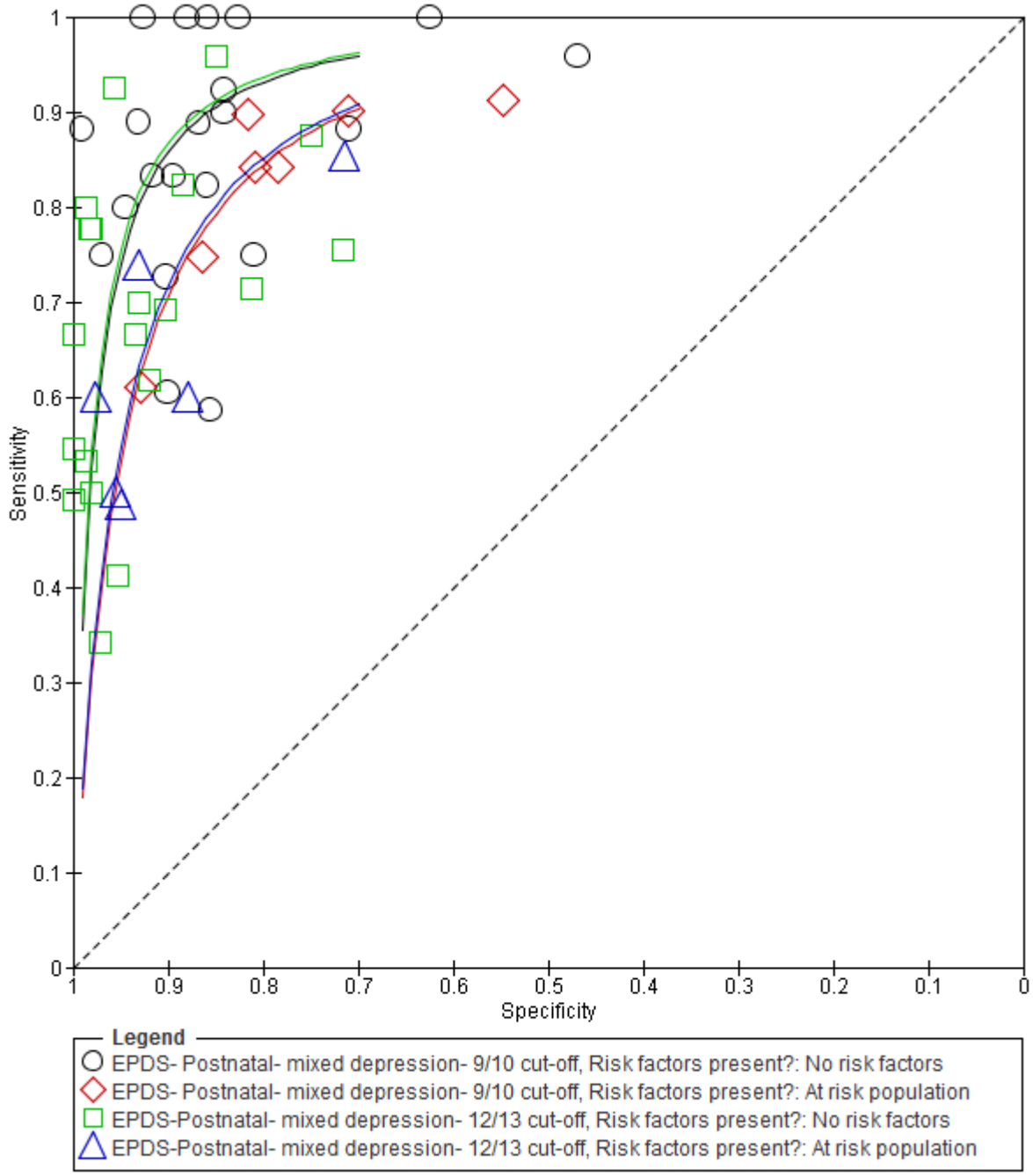


Figure 5: Summary of ROC curve for the EPDS administered in the postnatal period for mixed depression at different cut-off points, sub-grouped by population at risk of depression



Patient Health Questionnaire

The Patient Health Questionnaire (PHQ) developed out of the more detailed Primary Care Evaluation of Mental Disorders (PRIME-MD) (Spitzer et al., 1994). A nine-item depression module (PHQ-9) is often used in isolation, for example by GPs, and a two-item version (PHQ-2) has also been tested and found to have good sensitivity and specificity (Kroenke et al., 2003). The PHQ-9 has a cut-off of 10 and

the PHQ-2 follows the scoring format of the PHQ-9 (Likert scales) and has a recommended cut-off of 3 or 4.

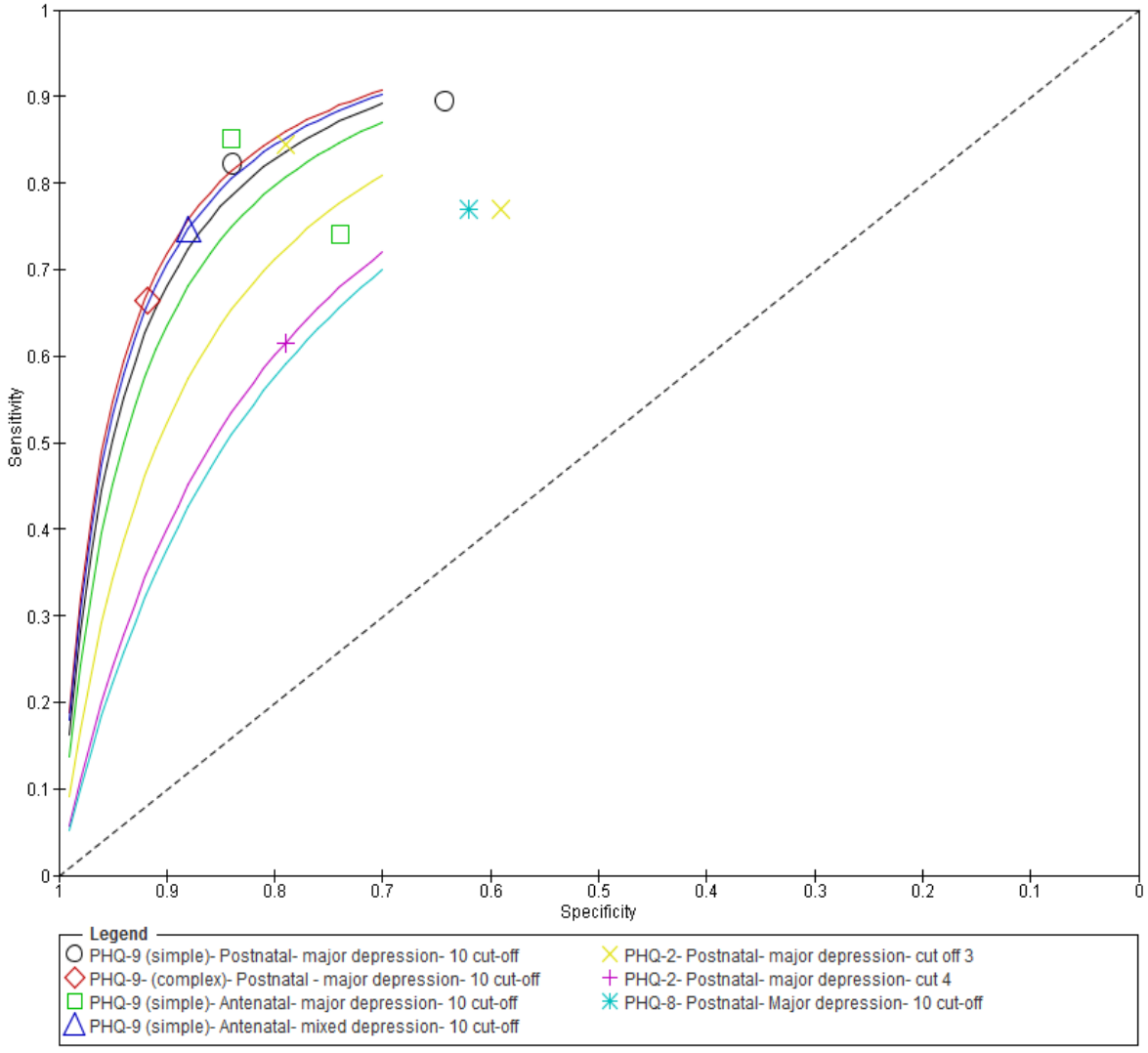
There were four studies investigating the PHQ in pregnancy and the postnatal period. A meta-analysis was not possible as there were insufficient data for each version of the PHQ at different timings and different types of diagnoses. Table 17 and Figure 6 summarise the sensitivity and specificity for PHQ items -2, -8 and -9 at different timings and diagnoses. See forest plots in Appendix 19 for individual data by study and the full methodological checklists in Appendix 17. The PHQ-2 had moderate to good sensitivity and low to moderate specificity at the cut-off 2/3, and moderate to good sensitivity and specificity at the higher cut-off 3/4 for detecting major depression in the postnatal period. In pregnancy the PHQ-9, at the cut-off 9/10 had good sensitivity and moderate to good specificity for detecting major and mixed depression. In the postnatal period, the simple version of the PHQ-9 had good to excellent sensitivity and moderate to good specificity. When the complex version of the PHQ-9 was used the sensitivity was lower, but the specificity higher.

Table 17: Evidence summary table for the PHQ (2-, 8- and -9 items)

Version **Cut-off** **Diagnosis**	**Timing**	**No of Participants (studies)**	**Sensitivity range (95% CI)**	**Specificity range** **(95% CI)**
PHQ-2 Cut-off 2/3 Major depression	Postnatal	719 (2)	0.84 (0.71-0.94) 0.77 (0.46-0.95)	0.79 (0.75-0.83) 0.59 (0.53-0.66)
PHQ-2 Cut-off 3/4 Major depression	Postnatal	213 (1)	0.63 (0.32-0.86)	0.79 (0.73-0.84)
PHQ-8 Cut-off 9/10 Major depression	Postnatal	213 (1)	0.77 (0.46-0.95)	0.62 (0.55-0.69)
PHQ-9 (simple scoring[1]) Cut-off 9/10 Major depression	Postnatal	605 (2)	0.89 (0.80-0.95) 0.82 (0.68-0.92)	0.65 (0.43-0.84) 0.84 (0.80-0.87)
PHQ-9 (simple[1]) Cut-off 9/10 Major depression	Pregnancy	814 (2)	0.74 (0.61-0.85) 0.85 (0.66-0.96)	0.73 (0.38-0.94) 0.84 (0.81-0.87)
PHQ-9 (complex scoring[2]) Cut-off 9/10 Major depression	Postnatal	506 (1)	0.67 (0.51-0.80)	0.92 (0.89-0.94)
PHQ-9 (simple[1]) Cut-off 9/10 Mixed depression	Pregnancy	745 (1)	0.75 (0.64-0.84)	0.88 (0.85-90)
Note. [1]Simple scoring: result is positive if sum of numbered responses is ≥10.				

[2]Complex scoring: result is positive if at least 5 symptoms are present, including symptom 1, symptom 2, or both, and each symptom present has a response score of 2 to 3, except for symptom 9, for which a response score of 1 to 3 was acceptable.

Figure 6: Summary of ROC curve for the PHQ (2-, 8- and 9-item versions) at different timings, diagnoses and cut-offs



Whooley questions

The 'Whooley questions' involve two brief focused questions that address mood and interest ('During the last month, have you often been bothered by feeling down, depressed or hopeless?' and 'During the last month have you often been bothered by having little interest or pleasure in doing things?'); studies indicate that these questions are as likely to be effective as more elaborate methods and are more compatible with routine use in busy primary and secondary care settings (Whooley et al., 1997). The questions are based on the 2-item PHQ-9 (see above), although in the Whooley version the questions are not scored but simply require a yes or no

answer. Arroll and colleagues (2005) developed an extension to these two questions by adding the following question: 'Is this something with which you would like help?'.

There were two studies which validated the Whooley questions in pregnancy and the postnatal period.

Table 18 and Figure 7 summarise the sensitivity and specificity for the Whooley questions at different timings and diagnoses. See forest plots in Appendix 19 for individual data by study and the full methodological checklists in Appendix 17One UK based study validated the two case-finding Whooley questions and also the addition of the third question about the need for help. In pregnancy the two case-finding questions had a sensitivity of 100%, however only moderate specificity for identifying mixed depression. Among women who screened positive in pregnancy, the additional 'help' question had excellent specificity but low sensitivity. The results for the two case-finding questions was similar in the postnatal period, however there was a lower sensitivity (39%) and higher specificity (100%) for the additional 'help' question.

Table 18: Evidence summary table for the Whooley questions

Tool version Diagnosis	Timing	No of Participants (studies)	Sensitivity range (95% CI)	Specificity range (95% CI)
Whooley questions Mixed depression	Postnatal	94 (1)	1.00 (0.81-1.0)	0.64 (0.53- 0.75)
Whooley questions Mixed depression	Pregnancy	126 (1)	1.00 (0.80-1.0)	0.68 (0.58-0.77)
Whooley questions (+ help question) Mixed depression	Postnatal	45 (1)	0.39 (0.17-0.64)	1.00 (0.87-1.0)
Whooley questions (+ help question) Mixed depression	Pregnancy	52 (1)	0.59 (0.33-0.82)	0.91 (0.77-0.98)
Whooley questions Major depression	Postnatal	506 (1)	1.00 (0.92-1.0)	0.44 (0.39-0.49)

Kessler-10

The Kessler-10 (Kessler et al., 2002) consists of ten self-report items based on a 4-week recall period. Participants respond to each item by rating the psychological distress experienced by them on a five point Likert scale. Each response is scored from 0 to 4 yielding a total score in the range of 0–40.

Three studies were found that assessed the Kessler-10 in pregnancy and the postnatal period; two during pregnancy and one in the postnatal period.

Table 19 summarises the sensitivity and specificity data. See forest plots in Appendix 19 for individual data by study and the full methodological checklists in Appendix 17.All studies were conducted in developing countries. One study demonstrated excellent sensitivity and good specificity in detecting major depression in pregnancy using a cut-off of 6, whilst another study reported only a moderate sensitivity and specificity. In the postnatal period, there was one study which found a good specificity but poor sensitivity using a cut-off of 6 to detect mixed depression, although the paper reported the optimum cut-off to be 12.

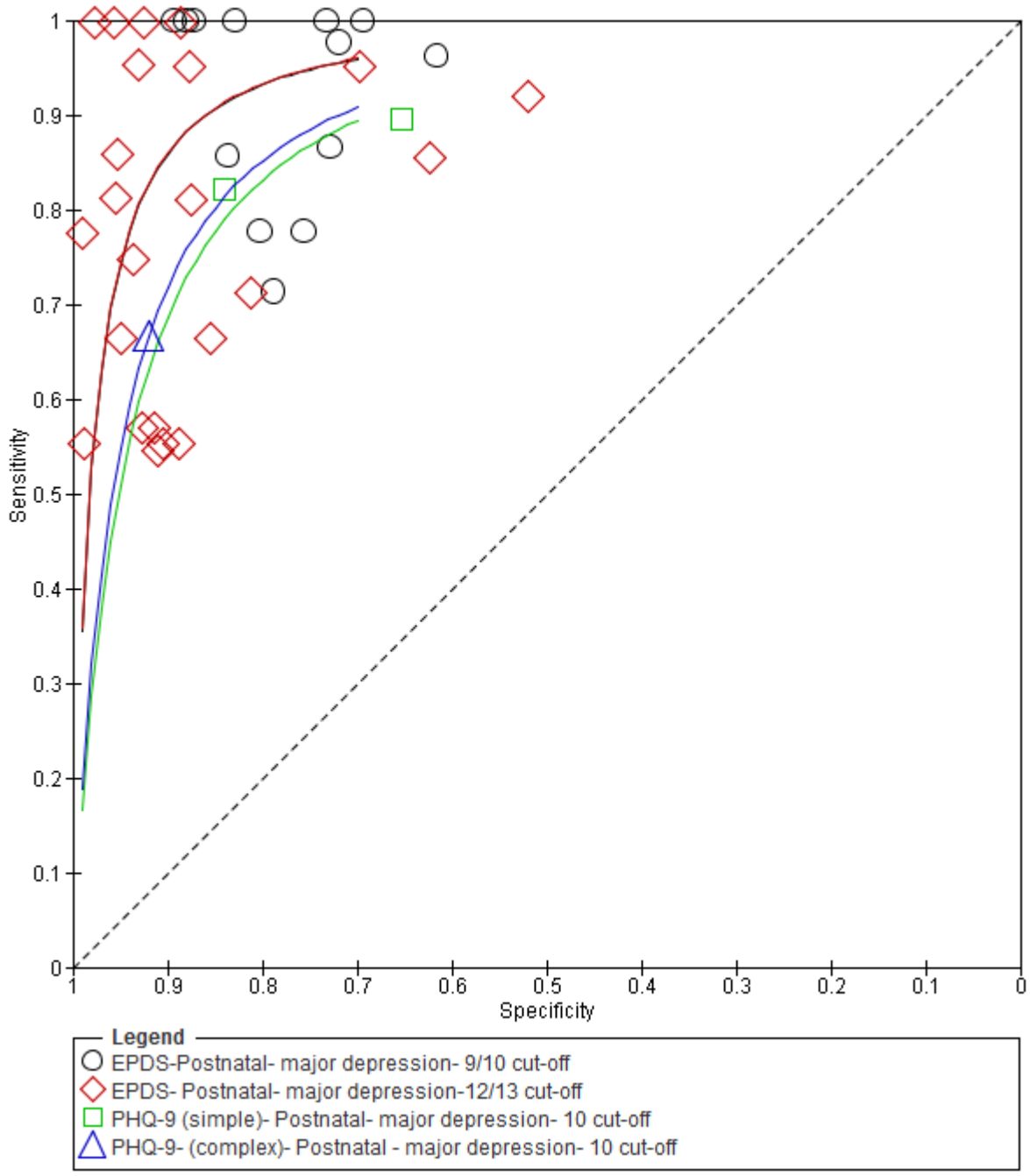
Table 19: Evidence summary table for the Kessler-10

Tool version **Diagnosis** **Cut-off**	**Timing**	**No of Participants (studies)**	**Sensitivity (95% CI)**	**Specificity (95% CI)**
Kessler-10 Major depression 6	Pregnancy	323 (2)	1.00 (0.88, 1.00) 0.75 (0.48, 0.93)	0.81 (0.74, 0.86) 0.54 (0.44, 0.63)
Kessler-10 Mixed depression 6	Postnatal	61 (1)	0.85 (0.66, 0.96)	0.41 (0.25, 0.59)

Comparison of different tools

It was only possible to make a comparison between the EPDS and PHQ-9 for detecting major depression in the postnatal period. Figure 7 presents a summary ROC curve comparing the EPDS and PHQ-9 in the postnatal period at different cut-off points.

Figure 7: Summary of ROC curve for the EPDS and PHQ- 9 for detecting major depression in the postnatal period at different cut-offs



Evaluating identification tools for anxiety

Overall the data on anxiety disorders was much more limited compared to depression therefore pooled analyses could not be conducted. When evaluating instruments, consideration was given to:

- The type of anxiety disorder that the gold standard diagnostic interview was used to classify; some studies evaluated the diagnostic accuracy of instruments for detecting 'anxiety disorders' as a broad term, however some specified the type of anxiety disorder (for example, PTSD)
- The timing at which the instrument was administered; in pregnancy or in the postnatal period.
- The cut-off point chosen to indicate a positive test; threshold effects can create a potential source of heterogeneity.

Edinburgh Postnatal Depression Scale

Three items (items 3, 4 and 5) from the full scale EPDS have been found to load onto an 'anxiety' factor known as the EPDS-3A in both pregnancy and the postnatal period and may be useful in detecting anxiety disorders (Matthey et al., 2008).

Of the eligible studies, there were two studies which evaluated the EPDS-3A for anxiety disorders (general anxiety disorder, panic disorder and OCD) and one which also included social phobia, specific phobia, and anxiety disorder not otherwise specified in their definition of anxiety disorders.

Table 20 summarises the sensitivity and specificity data for the EPDS at four different cut-off points in the postnatal period. One study found an optimum cut-off of 5/6 had only a moderate sensitivity but a good specificity, whereas the other found an optimum cut-off of 3/4 with only a moderate sensitivity and specificity. One study assessed the EPDS for detecting common mental health problems (depression and anxiety); at the optimal cut-off 3/4 they found moderate sensitivity and specificity.

Table 20: Evidence summary table for the EPDS for detecting anxiety

Tool version **Timing** **Diagnosis**	**Cut-off point**	**No of Participants (studies)**	**Sensitivity (95% CI)**	**Specificity (95% CI)**
EPDS-3 Postnatal Anxiety disorder	3/4	403 (2)	0.72 (0.47-0.90) 0.63 (0.49-0.76)	0.57 (0.50-0.63) 0.70 (0.61-0.79)
EPDS-3 Postnatal Anxiety disorder	4/5	403 (2)	0.67 (0.41-0.87) 0.47 (0.34-0.61)	0.73 (0.67-0.79) 0.90 (0.83-0.95)
EPDS-3 Postnatal Anxiety disorder	5/6	403 (2)	0.67 (0.41-0.87) 0.26 (0.16-0.40)	0.88 (0.83-0.92) 0.90 (0.83-0.95)
EPDS- full scale Pregnancy Anxiety disorder	8/9	200 (1)	0.80 (0.59-0.93)	0.68 (0.61-0.75)
EPDS-3 Pregnancy and postnatal	2/3	364 (1)	0.73 (0.64-0.81)	0.64 (0.58-0.70)

Anxiety and depression				
EPDS-3 Pregnancy and postnatal Anxiety and depression	3/4	364 (1)	0.70 (0.60-0.78)	0.73 (0.67-0.78)
EPDS-3 Pregnancy and postnatal Anxiety and depression	4/5	364 (1)	0.63 (0.54-0.72	0.81 (0.76-0.86)
EPDS-3 Pregnancy and postnatal Anxiety and depression	5/6	364 (1)	0.50 (0.41-0.60)	0.86 (0.81-0.90)

Kessler-10

Of the eligible studies there was one which assessed the Kessler-10 for identifying anxiety in pregnancy, which was explored for panic disorder, social anxiety and PTSD.

Table 21 summarises the sensitivity and specificity data for the Kessler-10 at the optimal cut-off points for the three anxiety disorders. The sensitivity and specificity estimates were inconsistent, and the confidence intervals were very wide for sensitivity measures.

Table 21: Evidence summary table for the Kessler-10 for detecting anxiety

Tool version **Timing** **Diagnosis**	**Cut-off point**	**No of Participants (studies)**	**Sensitivity (95% CI)**	**Specificity (95% CI)**
Kessler-10 Pregnancy Panic disorder	NR	129 (1)	0.50 (0.01, 0.99)	0.98 (0.93, 1.00)
Kessler-10 Pregnancy Social anxiety	NR	129 (1)	1.00 (0.03, 1.00)	0.75 (0.67, 0.82)
Kessler-10 Pregnancy PTSD	NR	129 (1)	0.50 (0.07, 0.93)	0.80 (0.72, 0.87)

5.3.5 Clinical evidence summary for case identification instruments for detecting mental health problems in pregnancy and the postnatal period

Identification of depression

Four brief case identification instruments were included in the review for detecting depression. The EPDS was the only tool where there was enough data to synthesise the results using meta-analysis and provide pooled summary estimates of sensitivity

and specificity. The GDG considered the diagnostic test accuracy results together with concerns about the methodological quality.

There were a substantial number of studies validating the EPDS in the postnatal period. For mixed depression sensitivity and specificity ranged from 34% to 100%, and from 47% to 100%, respectively. For major depression only, sensitivity ranged from 55% to 100% and specificity from 52% to 99%. When deciding an optimal cut-off point, the GDG considered the trade-off between sensitivity and specificity. Using the pooled estimates from the meta-analysis, the EPDS had good sensitivity and specificity for detecting major and minor depression at the lower cut-off 9/10. When increasing the cut-off to 12/13, the sensitivity decreased and the specificity increased; this would result in more women being missed but less being wrongly diagnosed.

There was substantial between-study heterogeneity found for almost all pooled estimates. This may have been due to differences in study design, population sampled, the timing of testing, different language version of the EPDS and the diagnostic criteria used. In addition, samples were conducted in a variety of clinical, community and research settings and drawn from women with different socioeconomic statuses, and from different countries with different cultural attitudes towards distress. The prevalence of depression also varied across studies and was over-represented in some. In order to address the heterogeneity, subgroups of interest were analysed separately for country (developed or developing), study design (cohort or case-control) and population (women with risk factors for depression or no risk factors for depression), however this had little impact on reducing the heterogeneity. Caution should therefore be taken when interpreting the results.

There were fewer studies validating the EPDS in pregnancy and there was a wide range of reported sensitivity and specificity measures across studies and substantial heterogeneity. Studies were conducted at different trimesters of pregnancy which may have been a possible source of heterogeneity, however subgroup analyses by trimester could not be conducted as there was insufficient data reported for each trimester. Given that the dataset had a number of problems, and no established cut-off point, the GDG did not feel it was sufficient to make a judgement about its utility in pregnancy.

There were two studies which evaluated the Whooley questions in the postnatal period, one a UK population validation study (Mann et al., 2012) which also evaluated its use in pregnancy. Both studies found the sensitivity to be 100%, suggesting the Whooley questions could provide as a simple approach to ruling out depression. However the specificity was a low and a substantial number of false-positives were found in both studies. These findings are similar to validation studies in the general population (Arroll et al., 2005). Mann and Gilbody (2012) did not find the additional question about the need for help had conclusive benefit, and resulted in poor discrimination between true-negative and false-negative cases which may

lead to an increased risk of depression being missed or lost to follow-up. However, the benefit of using a brief case-finding approach in clinical settings where routine perinatal care takes place is not necessarily to diagnose depression per se, but to reduce the number of women who need extensive assessment or evaluation with longer questionnaires such as the EPDS. Current NICE guidelines for depression (NICE, 2009a; NICE, 2009b) recommend the use of the two Whooley questions. The questions do not require additional resources (such as copies of a questionnaire), and the value lies in part in their brevity and the fact that they lend themselves to the use in both pregnancy and the postnatal period.

There was limited and insufficient evidence for the use of the Kessler-10 in pregnancy and the postnatal period. Like the EPDS, the PHQ, in particular the PHQ-9, also had good to excellent measures of sensitivity and specificity scores across a range of cut-offs and diagnoses, however it must be noted that there were substantially fewer studies validating the PHQ than the EPDS in this population and a pooled meta-analysis was not possible. When considering the administration of the EPDS and PHQ, the GDG favoured sensitivity over specificity (therefore a lower-cut-off) as appropriate, given that the role will be used in a group where the suspicion of depression had already been raised and for detecting women with subthreshold symptoms (both minor and major depression) rather than major depression only.

The GDG was conscious of the limited evidence base identified for instruments other than the EPDS in the reviews above. Case finding is most conveniently undertaken by healthcare professionals in regular contact with women, but they do not traditionally have training in mental health. The Whooley questions appear to offer a relatively quick and convenient way of case finding for healthcare professionals who are not specialists in mental health. The questions are suitable for a population-wide screen and would help to minimise unnecessary screening with longer tools for those who clearly do not meet depression criteria, by ruling these out. The EPDS or PHQ-9 appear to be suitable instruments for further assessment and have evidence for good sensitivity and specificity. Whilst, more timely to conduct, administration of the EPDS or PHQ-9 following a positive response to the Whooley questions may offer a way to decrease the number of false-negatives and allow the clinician to develop a clear idea of the nature of the clients problems.

Identification of anxiety disorders

Overall, there was limited evidence on the diagnostic accuracy of case identification instruments for detecting anxiety disorders in the perinatal period. Two studies reported on the validation of the three-item version of the EPDS, however demonstrated only moderate sensitivity and specificity at different optimum cut-offs in detecting anxiety disorders in the postnatal period. One study assessed the use of the full scale EPDS in detecting anxiety disorders in pregnancy, however found only moderate specificity. A further study reported on the use of the three-item version of the EPDS in detecting both anxiety disorders and depression during pregnancy and the postnatal period, however did not demonstrate good sensitivity and specificity at any of the different cut-offs. There was evidence from a single study for the use of

the Kessler-10 for detecting different types of anxiety disorders; however this tool did not demonstrate good sensitivity and specificity.

Given the limited evidence on the diagnostic accuracy of formal case identification tools for detecting anxiety disorders in pregnancy or the postnatal period and the recognition of the GDG of the significant impact these disorders have on both the woman and fetus, the GDG felt it better to draw on the more robust evidence base for case identification tools from other guidelines in non-pregnant population. This included the common mental health disorders guideline (NICE, 2011b) which recommends the use of the Generalized Anxiety Disorder scale – 2 items (GAD-2) (and the additional use of a question to elicit avoidance, if needed) to identify anxiety disorders. However the GDG felt it important that clinicians should also bear in mind that some changes in mental state and functioning are a normal part of the pregnancy and postnatal experience and should, therefore pay careful consideration to the context.

5.3.6 Health economic evidence

Systematic literature review

The systematic literature search identified one eligible UK study (Hewitt et al., 2009; Paulden et al., 2009) and one study conducted in New Zealand (Campbell et al., 2008) that assessed the cost effectiveness of case identification methods of mental health problems in women in the postnatal period. Both identified studies assessed the cost effectiveness of formal case identification tools for depression in the postnatal period. Details on the methods used for the systematic search of the economic literature are described in Chapter 3. References to included studies and evidence tables for all economic studies included in the guideline systematic literature review are presented in Appendix 21. Completed methodology checklists of the studies are provided in Appendix 20. Economic evidence profiles of studies considered during guideline development (that is studies that fully or partly met the applicability and quality criteria) are presented in Appendix 22, accompanying the respective GRADE clinical evidence profiles.

Paulden and colleagues (2009) evaluated the cost-utility of formal case identification methods for depression in the postnatal period compared with standard care for a hypothetical cohort of postnatal women managed in primary care. Hewitt and colleagues (2009) reported the same analysis as part of a HTA Assessment report. The authors used decision-analytic economic modelling to assess different case identification methods including EPDS with cut-off points ranging from 7 to 16; Beck Depression Inventory (BDI) cut-off point of 10; and also Whooley questions as part of the sensitivity analysis. Standard care was defined as opportunistic case finding. Case identification tools were administered 6 weeks after childbirth. In the base-case analysis mild and severe depression in the postnatal period were considered. Women that were identified with depression in the postnatal period were offered individual structured psychological therapy. The effectiveness data (that is, sensitivity and specificity) of the alternative formal identification methods were

derived from a bivariate meta-analysis. Resource use estimates were derived from various published sources and supplemented with authors assumptions where necessary; unit cost data were taken from national sources and other published literature. The time horizon of the analysis was 12 months and the perspective was that of NHS and personal social services (PSS). The study estimated costs associated with instrument administration, licence fees, subsequent treatment including health visitor, clinical psychologist, psychiatrist, GP, drug acquisition; and the costs associated with managing incorrect diagnosis. The measure of outcome for the economic analysis was the quality adjusted life year (QALY).

According to the model, the mean expected QALYs per woman was 0.846 to 0.847 for EPDS (cut-off points 16 to 8, respectively); was 0.847 for BDI (cut-off point 10); and 0.846 for standard care. The mean expected cost associated with the use of EPDS (cut-off points 16 to 8) was £74 to £215 per woman, respectively; with BDI (cut-off point 10) £122 per woman and with standard care it was £49 per woman in 2006/2007 prices. In the base-case analysis the identification strategies were ranked in terms of cost (from the least expensive to the most costly). The incremental cost effectiveness ratios (ICERs) were calculated for each successive alternatives (only after excluding dominated or extendedly dominated strategies). ICERs for all formal identification methods were above £40,000/QALY. The lowest ICER of £41,103/QALY was associated with EPDS cut-off point 16 (versus standard care). The ICERs for all other screening strategies ranged from £49,928/QALY (EPDS cut-off point 14 versus EPDS cut-off point 16) to £272,463/QALY (EPDS cut-off point 8 versus EPDS cut-off point 9). Probabilistic analysis indicated that at willingness to pay (WTP) of £20,000-£30,000/QALY the probability that standard care is cost effective was 0.877 to 0.587 (versus EPDS cut-off point 16). In the base-case analysis it was assumed that false positives would incur the costs of additional care (one CPN visit of 1 hour, three GP visits of 10 minutes each and four health visitor home visits of 45 minutes each) before being correctly diagnosed. However, assuming that false positives will be correctly diagnosed with a single GP consultation EPDS cut-off point 10 resulted in an ICER of £29,186/QALY when compared with standard care, which is just below NICE's upper cost-effectiveness threshold value of £30,000/QALY. Furthermore, using EPDS cut-off point 13 with confirmatory structured clinical interview resulted in an ICER of £33,776/QALY when compared with standard care; and using Whooley questions as an identification method resulted in an ICER of £46,538/QALY when compared with EPDS cut-off point 16. Also, when considering women only with severe depression in the postnatal period EPDS cut-off point 16 (versus standard care) resulted in an ICER of £23,195/QALY which is below NICE's upper cost-effectiveness threshold value of £30,000/QALY. Overall, the authors concluded that none of the case identification methods are cost effective for identifying depression in the postnatal period.

The analysis is directly applicable to this guideline review and the NICE reference case. This was UK-based study with QALYs as an outcome measure; however the utility values were not specific to women with depression in the postnatal period, due to lack of relevant data, but for the general population with depression treated

with antidepressant medication. The analysis assumed that positive response to the Whooley questions resulted in the provision of intensive psychological therapy and did not consider the possibility of further assessment. Also, a zero rate of false positives was assumed for standard care; however research by Mitchell and colleagues (2009) suggests that the false positive rate may be in the region of 15%. On the basis of the above, the GDG considered that the model structure did not adequately reflect the management of depression in the postnatal period in the UK. Consequently, the study was judged by the GDG to have potentially serious methodological limitations.

Campbell and colleagues (2008) evaluated the cost effectiveness and cost-utility of formal case identification programme compared with standard care in postnatal women attending Well Child Clinics in New Zealand. Formal case identification comprised three-question PHQ for depression in the postnatal period, administered at 6 weeks after childbirth by a GP or practice nurse, and again at 4 months after childbirth administered by a 'Well Child provider'. Standard care was defined as postnatal assessment using EPDS at core Well Child contacts at 6 weeks, 3 and 5 months, and other opportunistic contacts. Treatment of depression in the postnatal period comprised antidepressants and/or psychological therapy, or social support. This was a modelling study with effectiveness data (that is, sensitivity and specificity) of the alternative identification strategies derived from an observational study. The resource use estimates were based on national recommendations, international guidance, including the previous *Antenatal and Postnatal Mental Health* guideline (NICE, 2007a; NCCMH, 2007), other published sources, expert opinion and authors' assumptions; and the unit costs were obtained from national sources. The time horizon of the analysis was 12 months. The study estimated direct medical costs associated with screening and treatment including the provision of social support, psychological therapy and antidepressant medication; inpatient care, GP practice nurse, clinical psychologist, community counsellor and other prescriptions. The measure of outcome for the economic analysis was cases with depression detected and avoided in the postnatal period, and QALYs.

For the annual cohort of 56,635 women covered by the Well Child/Tamariki Ora programme formal case identification strategy resulted in a greater number of cases detected with depression in the postnatal period: 13,781 and 6,361 in intervention and standard care groups, respectively (difference of 7,420 cases); it also resulted in a greater number of cases of depression in the postnatal period that were resolved: 9,900 and 4,570 in intervention and standard care groups, respectively (difference of 5,330 cases). Intervention also resulted in a greater number of QALYs: 46,875 and 46,259 in intervention and standard care groups, respectively (difference of 616 QALYs). The costs in the study were measured in New Zealand dollars in 2006/2007 prices. The cost for the annul cohort of postnatal women over 12 months was NZ$3.9 million for intervention and NZ$1.7 million for standard care group, difference of NZ$2.1 million. The cost per additional case of depression in the postnatal period detected with the intervention compared with standard care was NZ$287; the cost per additional case of depression in the postnatal period resolved was NZ$400 and

the cost per QALY gained was NZ$3,461. The authors conducted extensive sensitivity analyses and the model was found to be most sensitive to the proportion of women that had depression that accessed and initiated appropriate treatment (that is, treatment uptake rate). Results suggest that a formal case identification programme is highly cost effective for depression in the postnatal period in New Zealand. The ICER of NZ$3,461/QALY converted to UK pounds using purchasing power parities (PPP) exchange rates and uplifted to 2013/2014 UK pounds using the UK Hospital and Community Health Services (HCHS) inflation index would be equivalent to £1,759/QALY, which is well below NICE's lower cost-effectiveness threshold value of £20,000/QALY.

Overall this analysis was judged by the GDG to be partially applicable to this guideline review and the NICE reference case. The study was conducted in New Zealand where the healthcare system is sufficiently similar to UK NHS. Many assumptions in the model were based on the previous *Antenatal and Postnatal Mental Health* guideline (NICE, 2007a; NCCMH, 2007) and *Depression* (NICE, 2009a; NCCMH, 2009), nevertheless effectiveness and resources use data were supplemented with expert opinion and authors' assumptions; and utility values used were for general population with depression treated with antidepressant medication. Also, the model unrealistically assumed that GPs correctly identify all women (that is, no false positives were associated with the GP assessment). As a result, the study was judged by the GDG to have potentially serious methodological limitations.

Economic modelling

Introduction: the objective of economic modelling

Existing UK-based economic evidence on case identification of depression in the postnatal period was limited to one study. Even though the study by Paulden and colleagues (2009) was judged to be directly applicable to the decision problem, it was characterised by potentially serious methodological limitations. The cost effectiveness of different case identification methods for depression in the postnatal period was considered by the GDG as an area with significant resource implications. Also, the clinical evidence in this area was judged to be sufficient and of adequate quality to inform economic modelling. Therefore, an economic model was constructed to assess the relative cost effectiveness of formal identification methods for women with depression in the postnatal period in the UK.

In constructing this model, the GDG was concerned to model an element of the case identification and assessment pathway. Specifically, the model was designed to assess the relative cost effectiveness between the use of a brief case identification tool followed by a more formal assessment method, the use of EPDS only, and standard care, defined as GP assessment.

It should be noted that the economic model focused on depression in the postnatal period because this was the only area with data of adequate quality to enable economic modelling.

Study population

The model was constructed for a hypothetical cohort of 1,000 postnatal women undergoing screening for depression.

Economic modelling methods

Interventions assessed

The choice of formal identification tools assessed in the economic analysis was determined after reviewing available relevant clinical data included in the guideline meta-analysis and the expert opinion of the GDG. Based on these, the following identification strategies were assessed in the economic analysis:

- EPDS only
- Whooley questions followed by EPDS
- Whooley questions followed by PHQ-9.

The identification strategies were compared with each other and also with standard care case identification. Standard care case identification refers to the routine clinical assessment that healthcare professionals would undertake to arrive at an informed and consensual diagnosis of depression in the postnatal period (without the formal use of a diagnostic instrument), and was defined as GP assessment.

Model structure

A decision-analytic model in the form of a decision-tree was constructed using Microsoft Office Excel 2013. The model structure was based on the model developed by Paulden and colleagues (2009). According to the model, hypothetical cohorts of 1,000 postnatal women managed in the primary care were initiated on one of the case identification strategies 6 weeks after childbirth. Depending on whether women undertaking the test did or did not have depression and the outcome of the identification test, four groups of women were formed: true positive, true negative, false positive and false negative. All positive cases were assumed to undergo formal assessment that according to the GDG expert opinion in clinical practice would be performed by health visitors. It has to be noted that formal assessment of positive cases by health visitors was considered only in terms of costs since no studies could be identified that reported how the use of formal case identification affected the subsequent assessment by a clinician.

Each of the four groups was assigned to a care pathway and followed up until the model endpoint at 1 year after childbirth. Women who were found to be true positive for depression were assumed to receive one of the following treatment options, in proportions reflecting severity of depression in the postnatal period: women with sub-threshold/mild to moderate depression were assumed to receive

facilitated guided self-help (72%) and women with moderate to severe depression were assumed to receive high intensity psychological therapy (20%) and pharmacological treatment (8%). Based on the GDG expert opinion high-intensity interventions consisted of CBT or interpersonal psychotherapy (IPT) (16 sessions); pharmacological treatment consisted of sertraline for 8 weeks plus 6 months' maintenance. Women who were found to be false positive for depression received the same treatments in the same proportions as described for those who were found to be true positive, but were assumed to stop treatment earlier, and according to the GDG's estimate consumed only 20% of the healthcare resources (and consequently incurred 20% of the respective costs).

Women who were found to be false negative could get better on their own without any treatment (spontaneous recovery), in which case they were assumed to incur only health and social care costs until that point (that is, approximately 3 months after childbirth). However, if women did not get better on their own they were assumed to have one GP visit halfway through the follow-up period during which time the woman's depression could be detected and treatment would be offered in the same proportions as described for those women who were found to be true positive. On the other hand, if women were not detected by their GP during the follow-up they were assumed to continue to incur health and social care costs until the model endpoint. Women who were found to be true negative were assumed to receive no treatment and incur no health or social care costs. Owing to lack of relevant data, only first-line treatments were considered and relapse was not modelled. A schematic diagram of the case identification model is presented in Figure 8. Figure 9 and Figure 10 present the pathways for true positives and for false negatives, respectively.

Costs and outcomes considered in the analysis

The economic analysis adopted the perspective of the NHS and PSS, as recommended by NICE (NICE, 2012a). Therefore, only direct health and social care costs were considered in the model. Costs included identification costs (GP time or health visitor time), assessment costs (health visitor time), treatment costs for women identified as having depression in the postnatal period (facilitated guided self-help, high intensity psychological therapy and pharmacological treatment), and extra health and social care costs for those women that were not identified by one of the alternative strategies, or that were identified but did not respond to treatment. Health and social care costs included costs associated with the care of infants too. The measure of outcome was the QALY.

Clinical input parameters to the economic model

Table 22 reports the values of all input parameters, including clinical inputs that were utilised in the economic model. The prevalence of depression in the postnatal period was derived from a UK-based study conducted by Sharp and colleagues (2010). This was a pragmatic two-arm RCT that evaluated the clinical effectiveness of antidepressant treatment for women with depression in the postnatal period compared with general supportive care. The overall prevalence of depression in the

postnatal period among study participants (n=4,173) was 8.7%, based on a completed screening questionnaire (n=4,158) or GP/health visitor referral (n=15). Based on the Clinical Interview Schedule-Revised (CIS-R) scores it was estimated that at baseline 20% of women had mild depression, 59% moderate and 22% severe. According to the GDG expert opinion 10% of women presenting with moderate symptoms would tend towards the severe spectrum of the disorder. Consequently, in the economic model it was assumed that 28% of women would experience moderate to severe depression and the remaining 72% mild to moderate depression.

Figure 8: Schematic diagram of decision-tree constructed for case identification and assessment for women with depression in the postnatal period

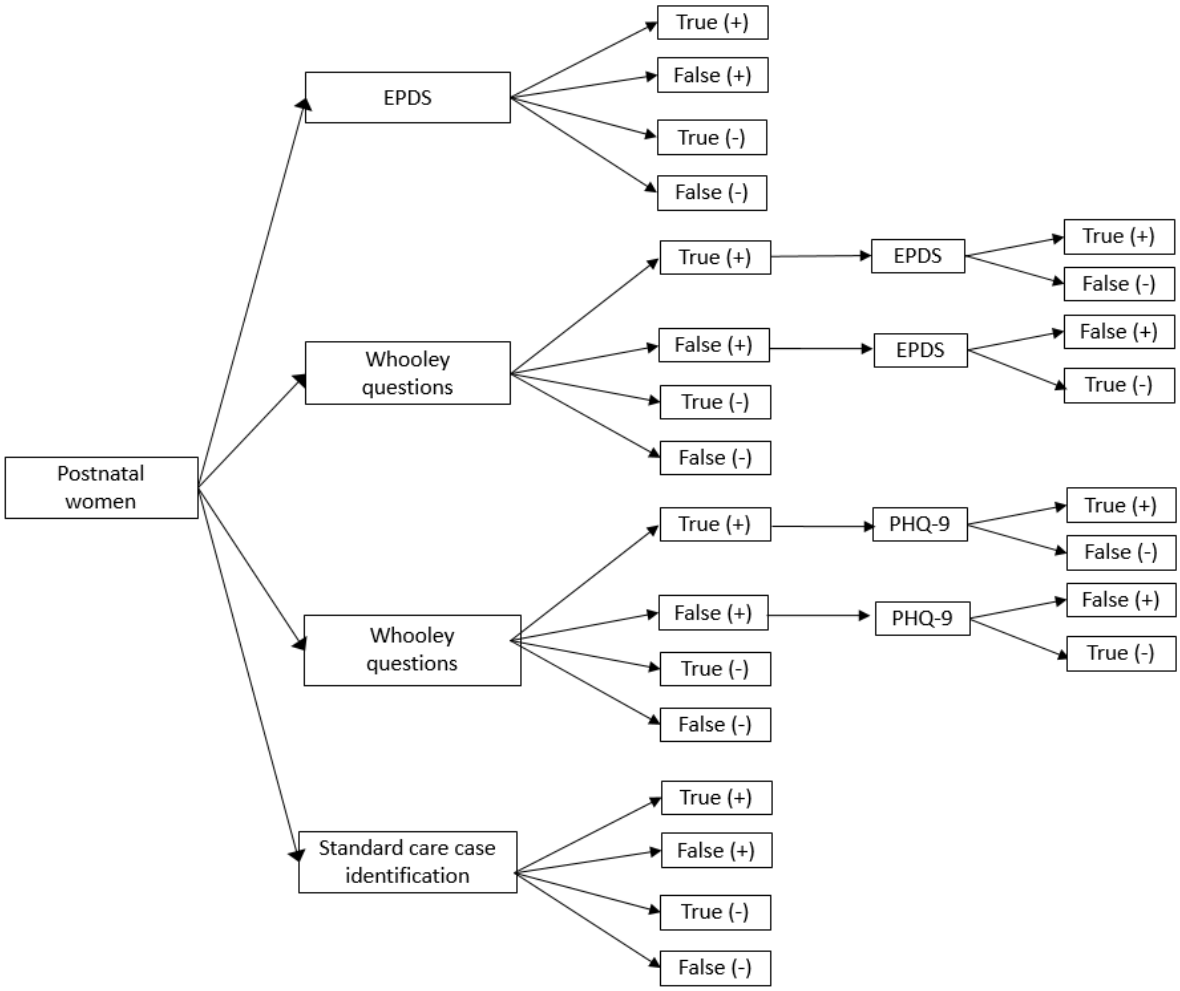


Figure 9: Pathway for true positives and false positives

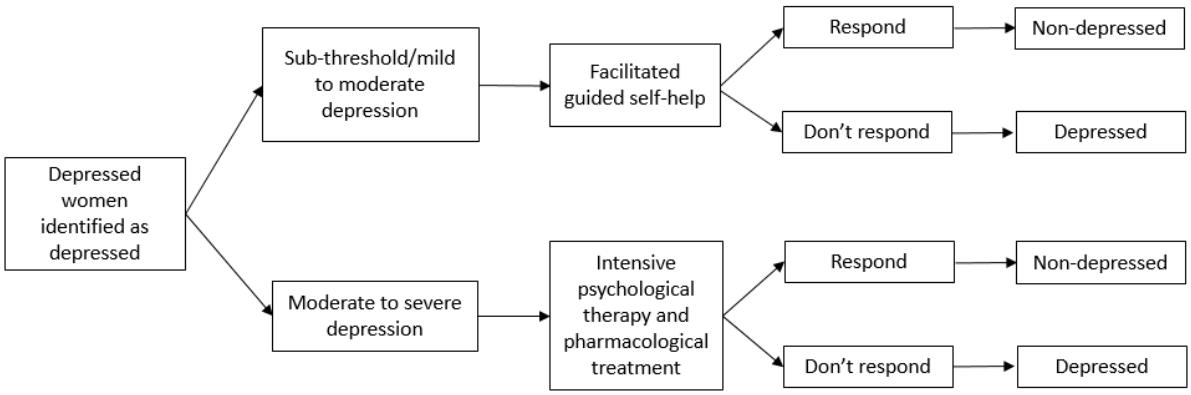
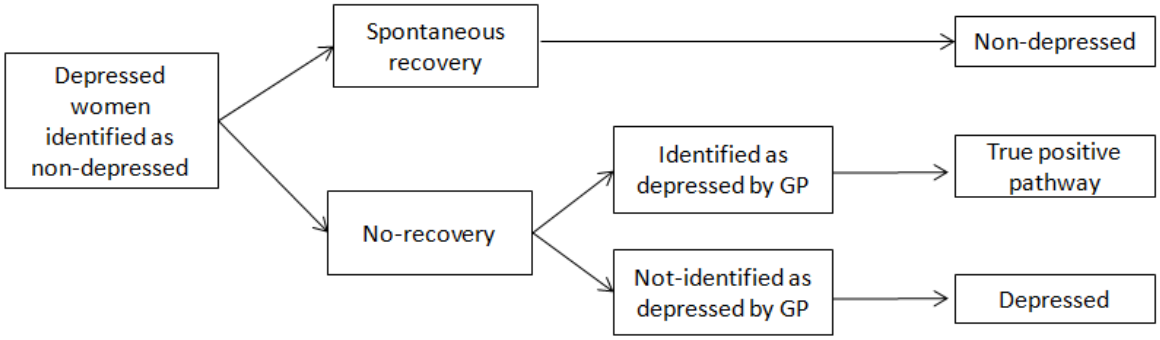


Figure 10: Pathway for false negatives



Clinical input parameters included the sensitivity and specificity of identification methods (standard care case identification, EPDS, PHQ-9 and Whooley questions). Sensitivity and specificity of the formal case identification methods were obtained from guideline meta-analysis. Sensitivity and specificity of:

- EPDS was for combined sub-threshold/mild and severe depression in the postnatal period; and a cut-off point of 9/10 was used
- PHQ-9 was for combined sub-threshold/mild and severe depression in the postnatal period; and a cut-off point of 10 was used
- Whooley questions was for combined sub-threshold/mild and severe depression in the postnatal period.

The GDG expressed their wish to focus on sub-threshold/mild to severe depression in the postnatal period hence in the model the cut-off of 9/10 was used for the EPDS and 10 for PHQ-9. No studies that met clinical review inclusion criteria and reported sensitivity and specificity for PHQ-9 administered in the postnatal period were identified; however the GDG judged that antenatal data should apply to the postnatal period as well. It should also be noted that most validation data available were for EPDS. Sensitivity and specificity for the PHQ-9 and Whooley questions were based on single studies. Also, because of a lack of relevant data, the model

assumed that sensitivity and specificity of the Whooley questions and any subsequent tests (that is, EPDS or PHQ-9) were independent of each other.

No studies were found that reported sensitivity and specificity for standard care case identification (that is, GP assessment) for the study population. Mitchell and colleagues (2009) conducted a meta-analysis of 118 studies that assessed the accuracy of diagnoses of depression by GPs. In their analysis 50,371 participants were pooled across 41 studies and examined. From these studies, the weighted sensitivity and specificity associated with GP assessment was 50.1% and 81.3%, respectively. These estimates were utilised in the economic model to approximate sensitivity and specificity associated with standard care case identification.

Regarding treatment, the response rate associated with facilitated guided self-help was obtained from a meta-analysis conducted for this guideline that included three RCTs (MILGROM2011A [Milgrom et al, 2011], OMAHEN2013A [O'Mahen et al., 2013a], OMAHEN2013C [O'Mahen et al., 2013c]) and intensive psychological therapy from six RCTs (AMMERMAN2013A/2013B [Ammerman et al., 2013a; Ammerman et al., 2013b], BURNS2013/PEARSON2013B [Burns et al., 2013; Pearson et al., 2013b], COOPER2003 [Cooper et al., 2003]/MURRAY2003 [Murray et al., 2003], GROTE2009, OHARA2000 [O'Hara et al., 2000], RAHMAN2008 [Rahman et al., 2008]). Women given pharmacological treatment were assumed to respond at the same rate as women treated with intensive psychological therapy.

In the model it was assumed that women who were found to be false negative could get better on their own without any treatment (spontaneous recovery). In the review by Dennis and colleagues (2009) it is reported that in trials of treatment for depression in the postnatal period spontaneous recovery rates in control groups range between 25-40%. In the analysis, the midpoint of 33% was used to approximate a proportion of women with a false negative result who would spontaneously enter remission; the majority of women who spontaneously improve on their own do so approximately by 3 months after childbirth (Royal College of Psychiatrists' Public Education Editorial Board, 2014). The reported spontaneous recovery rate of 33% is fully consistent with standard care arms of guideline meta-analyses (that is, the absolute risk of non-improvement is 67% implying the spontaneous recovery rate of 33%).

Also, a proportion of women with false negative result and who do not improve on their own could be detected by their GP during the follow-up. In the model it was assumed that these women would have one GP consultation halfway through the follow-up during which depression could be detected. No studies were identified that reported the probability of GPs detecting depression in the postnatal period during the follow-up. Kessler and colleagues (2002) conducted a study aiming to determine the probability of GPs diagnosing depression or anxiety during the follow-up given that it was not diagnosed during the initial consultation. The authors followed up consecutive attenders at a general practice in north Bristol in 1997. It was found that of the participants who had not received a diagnosis during

the initial consultation, 41% received a diagnosis during the 3 years' follow-up. Based on the above it was estimated that approximately 8% of cases would be detected by a follow-up consultation at 6 months.

Resource use and cost data

Costs associated with the case identification strategies were calculated by combining resource use estimates (that is, GP or health visitor time) with respective national unit costs (Curtis, 2013). According to the studies included in the guideline meta-analysis, use of EPDS and PHQ-9 requires approximately 15 minutes for each (that is, 10 minutes administration and 5 minutes scoring), and administration of Whooley questions requires approximately 1 minute; whereas based on the GDG expert opinion it was estimated that routine case identification required on average one GP consultation that would last approximately 11.7 minutes (Curtis, 2013). Moreover, according to the GDG expert opinion, formal case identification would be followed by an assessment that in clinical practice would be done by a health visitor and would last approximately an hour.

Costs of psychological treatments were estimated using estimates in the studies that were included in the guideline meta-analysis; where necessary these were supplemented by the GDG expert opinion. According to the GDG expert opinion, facilitated guided self-help would be provided with support by psychological wellbeing practitioners trained in the perinatal issues (on the Agenda for Change band 5 salary scale); a mean of seven (range, six to eight) face-to-face support sessions each lasting approximately 25 minutes would be required. The unit cost for psychological wellbeing practitioner was not available. The unit cost was approximated using the unit cost reported by Curtis (2013) for a mental health nurse of £74 per hour. This was based on the mean full-time equivalent basic salary for Agenda for Change band 5 of the July 2012-June 2013 NHS staff earnings estimates for qualified nurses. Also, the cost of guided self-help manual (that is, *Overcoming Depression: A Books on Prescription Title*) was estimated to be £9.09 (amazon.co.uk).

In studies included in the guideline meta-analysis of intensive psychological therapies, treatment comprised of 9-21 individual sessions, however the GDG judged that in clinical practice women with moderate to severe depression in the postnatal period would receive approximately 16 sessions. The unit cost of intensive psychological therapy was estimated using the unit cost for CBT obtained from Curtis (2013). The unit cost was based on a full-time equivalent basic salary of the July 2012-June 2013 NHS staff earnings estimates for a specialty doctor (midpoint), clinical psychologist (band 8) and mental health nurse (band 5).

Also, according to the GDG expert opinion women receiving facilitated guided self-help and intensive psychological therapy would require additional care that would comprise of 3 GP consultations. The unit costs of a GP consultation (£45) was taken from the latest PSSRU estimates (Curtis, 2013).

According to the GDG's expert opinion, approximately 25 to 30% of women with moderate to severe depression in the postnatal period would be offered antidepressant treatment. In the analysis, the midpoint of 28% was used to approximate a proportion of women who would be offered antidepressant treatment. The most common antidepressant prescribed would be sertraline. Sertraline acquisition cost was obtained from the Electronic Drug Tariff (NHS, Business Service Authority & NHS Prescription Services, 2014). The daily dosage of the drug was informed by the GDG expert opinion (that is, 50 mg per day). For women with moderate to severe depression in the postnatal period who were taking sertraline, the total cost of the drug was calculated over the 8 weeks of initial therapy plus 6 months' maintenance. Based on the GDG expert opinion all women with moderate to severe depression who receive antidepressant treatment would be actively monitored either in primary or secondary care during the initial treatment period. It was assumed that 15% of women over initial therapy of 8 weeks would have, on average, two consultant psychiatrist visits (the first consultation lasting 30 minutes and the second consultation 15 minutes); the remainder of the visits for these women would be with a GP. The rest of the women managed with antidepressants were assumed to be managed in primary care only and would require a mean of four GP consultations. The unit costs of a GP consultation (£45) and a mental health outpatient consultation with consultant psychiatrist (£273) were both taken from the latest PSSRU estimates (Curtis, 2013).

Women who were falsely detected as having depression in the postnatal period were assumed to incur 20% of the treatment cost of a true positive woman, according to the GDG's estimate. Women identified as false negative (that is, women having depression in the postnatal period but not identified by the methods assessed in the model), as well as women not responding to treatment, were assumed to incur health and social care costs as described by Petrou and colleagues (2002). Petrou and colleagues (2002) estimated the economic costs of depression in the postnatal period in a geographically defined cohort of women at high risk of developing the condition. Health and social care costs were estimated based on 206 women recruited from antenatal clinics and their babies. The study estimated costs associated with community care, day care services, hospital outpatient attendances, hospital inpatient admissions, and paediatric and child care services. Since health and social care costs reported by Petrou and colleagues (2002) included paediatric and child care services this partially enabled incorporation of costs associated with infant care into this economic analysis.

In the model it was assumed that all postnatal women, whether depressed or non-depressed, consumed the same amount of healthcare resources during the first 6 weeks after childbirth. As a result, these costs were assumed to be common for all strategies being evaluated and so were not considered in the analysis. Standard postnatal care costs were omitted from the analysis, because they were common to all options being assessed. Also, other costs to women and family, such as personal expenses and productivity losses were excluded as they were beyond the scope of the analysis. Intangible costs (negative impact of the woman's depression on her

child's cognitive and emotional development as well as distress to the family) were also not estimated, but they should be taken into account when interpreting the results.

All costs were expressed in 2013 prices. Discounting of costs and outcomes was not necessary since the time horizon of the analysis was 1 year.

Utility data and estimation of QALYs

To express outcomes in the form of QALYs, the health states of the economic model needed to be linked to appropriate utility scores. Utility scores represent the HRQoL associated with specific health states on a scale from 0 (death) to 1 (perfect health); they are estimated using preference-based measures that capture people's preferences on the HRQoL experienced in the health states under consideration. The systematic search of the literature did not identify any studies that reported utility scores for specific health states associated with depression in the postnatal period. As a result these were approximated using utility scores reported by Sapin and colleagues (2004) for the general population with depression.

The study by Sapin and colleagues (2004) was based on a multicentre, prospective cohort of service users (n=250) with a new episode of major depressive disorder recruited in the French primary care setting assessed at 8 weeks' follow-up. European Quality of Life – 5 Dimensions (EQ-5D) utility scores were stratified according to depression severity (defined by Clinical Global Impressions [CGI] Severity scores), and by clinical response (defined by Montgomery-Åsberg Depression Rating Scale [MADRS] scores) at follow-up. Based on the GDG expert opinion utility scores for 'sub-threshold/mild to moderate' depression were approximated using utility scores for 'slightly/moderately ill', for 'moderate to severe' depression utility scores for 'markedly ill' were used; 'no depression' health state was approximated using utility scores for 'first signs' depression (the value of which was also very similar to utility scores for 'responder remitters').

In the model women identified as true negatives were assigned utility score associated with 'no depression' health state until the model endpoint. No studies were identified that assessed the impact of false positive diagnosis in the study population. According to the GDG expert opinion, it was assumed that a false positive diagnosis would result in a reduction of ~2% in HRQoL (that is, the utility weight for women with false positive diagnosis would be 2% lower than the utility weight for 'no depression'). Women who received treatment and responded (that is, true positives and women detected by their GP during the follow-up) were assumed to experience a linear improvement in their HRQoL from the initiation of treatment until the end of treatment; and then remained in the 'no depression' health state until the model endpoint. Similarly, women who had a spontaneous recovery were assumed to experience a linear improvement in HRQoL over the 3 months and then remained in the 'no depression' health state until the model endpoint. Women who did not respond to treatment or were not detected by their GPs during the follow-up were assumed to remain at baseline utility (that is, they experienced HRQoL

associated with either 'sub-threshold/mild to moderate' depression or 'moderate to severe' depression) until the model endpoint.

Table 22 reports the values of all input parameters utilised in the economic model, and provides details on the sources of data and methods that were used in the estimation of input parameters.

Table 22: Input parameters utilised in the economic model of formal case identification methods for women with depression in the postnatal period

Input parameter	Deterministic value	Source of data– comments
Prevalence of depression in the postnatal period	8.7%	Sharp and colleagues (2010)
Severity of depression in the postnatal period: Sub-threshold/mild to moderate Moderate to severe	 72% 28%	Sharp and colleagues (2010); GDG expert opinion
Spontaneous recovery rate	33%	Dennis and colleagues (2009)
Sensitivity of identification methods: Whooley questions EPDS (cut-off 9-10) PHQ-9 (cut-off 10) Standard care case identification	 1.00 (0.81; 1.00) 0.83 (0.81; 0.86) 0.75 (0.64; 0.84) 0.50	Guideline meta-analysis; sensitivity and specificity are for combined sub-threshold and severe depression in the postnatal period Mitchell and colleagues (2009)
Specificity of identification methods: Whooley questions EPDS PHQ-9 Standard care case identification	 0.64 (0.53; 0.75) 0.85 (0.84; 0.86) 0.88 (0.85; 0.90) 0.81	Guideline meta-analysis; sensitivity and specificity are for combined sub-threshold and severe depression in the postnatal period Mitchell and colleagues (2009)
Tool administration time: Whooley questions EPDS PHQ-9 Standard care case identification	 1 minute 15 minutes 15 minutes 11.7 minutes (1 GP consultation)	Guideline meta-analysis The GDG expert opinion; Curtis (2013)

Relative risk of no improvement for: Facilitated guided self-help Intensive psychological therapy	 0.73 0.48	Guideline meta-analysis
Absolute risk of no improvement: Standard care (sub-threshold/mild to moderate depression) Standard care (moderate to severe depression)	 0.67 0.65	Guideline meta-analysis (standard care arms of guideline meta-analysis)
Utilities: No depression Sub-threshold/mild to moderate depression Moderate to severe depression Reduction in utility due to false (+) diagnosis	 0.86 0.74 0.44 2%	Sapin and colleagues (2004); data refer to the general patient population with depression The GDG expert opinion
Cost of facilitated guided self-help and additional care:	£359.92	Based on seven telephone-based support sessions (25 minutes per session) provided by psychological wellbeing practitioner (Band 5) trained in perinatal issues; plus guided self-help manual costing £9.09 (*Overcoming Depression: A 'Books on Prescription' Title: A Self-help Guide Using Cognitive Behavioral Techniques* by Paul Gilbert). According to the GDG expert opinion additional care would comprise three GP consultations. Unit cost of psychological wellbeing practitioner unavailable; unit cost approximated using unit cost of mental health nurse (Band 5) £74 per hour; unit cost of GP visit lasting 11.7 minutes, £45 (Curtis, 2013)
Cost of intensive psychological therapy and additional care:	£1,591.00	Intensive psychological therapy was estimated to consist of 16 sessions with each session lasting 55 minutes. According to the GDG expert opinion, additional care would comprise three GP consultations. Unit cost of psychological therapy per session £91; unit cost of GP visit lasting 11.7 minutes, £45 (Curtis, 2013)
Cost of pharmacological treatment and additional care:	£214.98	Based on pharmacological treatment with sertraline for 8 weeks plus 6 months' maintenance. Unit cost of sertraline £2.09 per 28, 50 mg tablets (NHS Business Services Authority & NHS Prescription Services, 2014). Fifteeen percent of women would have two consultations with consultant psychiatrist, lasting 30 minutes and 15 minutes, respectively, and two consultations with GP. The remainder 85% percent of women would have 4 GP consultations. Unit cost of consultant psychiatrist per patient-related hour £273; unit cost of GP visit lasting 11.7 minutes, £45 (Curtis, 2013)

Weekly health and social care cost incurred by women with depression in the postnatal period	£8.21	Petrou and colleagues (2002); Health and social care costs were applied to women that were false (-) following case identification; and also to women who did not respond to treatment. Costs reported were uplifted to 2013 UK pounds using UK HCHS inflation index.

Data analysis and presentation of the results

In order to take into account the uncertainty characterising the model input parameters sensitivity analysis was undertaken to investigate the robustness of the results under the uncertainty characterising some of the input parameters and the use of different assumptions in the estimation of the cost effectiveness of case identification methods for depression in the postnatal period. One-way and two-way sensitivity analyses explored the impact of the following factors and scenarios on the results and conclusions of the analysis:

- changes in a range of epidemiological inputs including prevalence of depression in the postnatal period (varying from 3 to 20%), and the proportion of women with moderate to severe depression (varying from 10 to 50%)
- the uncertainty characterising the sensitivity and specificity of the identification methods (estimates were varied by ± 5-20%). Furthermore, two-way sensitivity analyses on sensitivity and specificity were also performed to further investigate uncertainty around those parameters. A simultaneous change of ± 5-20% in those parameters was tested.
- changes in the relative risk estimates associated with facilitated guided self-help and intensive psychological therapy (estimates were varied by ± 10-20%).
- changes in the consultation time necessary for the performance of the EPDS and PHQ-9; time was varied from 5 minutes to 20 minutes.
- costs associated with false positive cases were varied from 10 to 50% of costs associated with true positives.
- the uncertainty characterising treatment costs (estimates were varied by ± 50%).
- current standard care case identification being done by a health visitor rather than a GP.
- assessment following formal case identification being done by a GP rather than a health visitor.

Moreover, threshold sensitivity analyses were conducted to explore the magnitude of change in base-case values for the conclusions of the cost-utility analysis to be reversed.

Probabilistic sensitivity analysis was not possible due to limitations in the data (that is, it was not possible to model interaction between sensitivity and specificity associated with Whooley questions or PHQ-9 since diagnostic characteristics for these tools were derived from single studies).

Validation of the economic model

The economic model (including the conceptual model and the excel spreadsheet) was developed by the health economist working on this guideline and checked by a

second modeller not working on the guideline. The model was tested for logical consistency by setting input parameters to null and extreme values and examining whether results changed in the expected direction. The results were discussed with the GDG for their plausibility.

Results

Full results of the base-case analysis are presented in Table 23. According to the analysis, accounting for both identification and treatment costs, identification of depression in the postnatal period using Whooley questions followed by PHQ-9 was estimated to be the most cost-effective case identification strategy. Even though Whooley questions followed by EPDS resulted in the highest number of QALYs among all case identification options, when compared with Whooley questions followed by PHQ-9, it led to a small incremental health gain of 0.113 QALYs at an additional cost of £5,163 (results per 1,000 women), resulting in an ICER of Whooley followed by EPDS versus Whooley followed by PHQ-9 of £45,593/QALY. This latter value is well above NICE's cost-effectiveness threshold value of £20,000-£30,000/QALY. All other options (namely EPDS only and standard care case identification) were dominated (that is, results in higher costs and lower QALYs) by strategies utilising Whooley questions.

Table 23: Mean costs and QALYs for each identification option for women with depression in the postnatal period assessed in the economic analysis – results for a hypothetical cohort of 1,000 women

Identification strategy	Mean total QALYs	Mean total costs	Incremental QALYs	Incremental costs	Cost effectiveness
Whooley questions followed by EPDS	752.09	£80,517	0.113	£5,163	ICER of Whooley questions followed by EPDS versus Whooley questions followed by PHQ-9: £45,593/QALY
Whooley questions followed by PHQ-9	751.98	£75,354	-	-	
EPDS only	750.76	£106,351	-1.334	£25,834	Dominated
Standard care case identification	749.16	£111,269	-1.597	£4,918	Dominated

One-way sensitivity analyses showed that only if the prevalence of depression in the postnatal period was approximately 20% the model's conclusions would change (that is, Whooley questions followed by EPDS would be the preferred case identification strategy with cost per QALY of £20,000). As the proportion of women with moderate to severe depression in the postnatal period was varied from 10 to 50% the conclusions of the analysis did not change; however as the proportion fell below 7% Whooley questions followed by PHQ-9 became the dominant case identification strategy (that is, it resulted in lowest costs and the highest number of QALYs among all strategies assessed in the analysis).

Model's conclusions were found to be sensitive to the values of sensitivity and specificity for PHQ-9 and EPDS. As specificity for PHQ-9 improved by approximately 5% (from the base-case value) Whooley questions followed by PHQ-9 became the dominant case identification strategy and when it deteriorated by approximately 5% Whooley questions followed by EPDS became the dominant option. Similarly, changes in the sensitivity or specificity for EPDS (changes of approximately ± 10%) reversed the above conclusions. The conclusions were not affected by changes in the sensitivity or specificity for Whooley questions. A two-way sensitivity analysis showed comparable results (that is, the model was sensitive to small simultaneous changes in the estimates of sensitivity and specificity for formal case identification methods).

The model was also found to be sensitive to the changes in the consultation time necessary for the performance of the EPDS. When EPDS administration time was reduced to 9 minutes only, Whooley questions followed by EPDS became the preferred identification strategy with an ICER of £20,000/QALY (when compared with Whooley questions followed by PHQ-9). On the contrary, the results were not affected by changes in the relative risk of no response of each of the two treatments considered; changes in the costs associated with false positives; changes in treatment costs; assuming that assessment following formal case identification was done by GP rather than health visitor); or that standard care identification was performed by a health visitor (rather than by GP).

Threshold sensitivity analyses showed that the results were sensitive to the diagnostic characteristics of formal case identification tools and also consultation time require to administer case identification tool. Full results of threshold sensitivity analyses are provided in Table 24.

Table 24: Results of threshold sensitivity analyses

Parameter	**Values that resulted in** **Whooley questions followed by EPDS the preferred strategy (ICER £20,000/QALY)**
Sensitivity for: EPDS PHQ-9 Whooley	 0.95 0.63 -
Specificity for: EPDS PHQ-9 Whooley	 0.87 0.86 0.85
Relative risk of no improvement associated with treatments: Facilitated guided self-help Intensive psychological therapy	 - 0.77 0.93
Prevalence of depression in the postnatal period	19.04%

Consultation time required to administer case identification tool: EPDS PHQ-9	 9 minutes 21 minutes

Discussion and limitations of the economic analysis

The results of the economic analysis suggest that the use of a formal case identification strategy that utilises a combination of Whooley questions and PHQ-9 is a cost-effective option. This finding is attributable to the fact that this strategy rules out a greater number of costly false positives and false negatives (compared with other strategies), combined with the fact that they can be easily and quickly performed by health visitors, resulting in relatively low intervention costs.

Although the data pertaining to the diagnostic characteristics associated with formal case identification tools were limited, extensive deterministic sensitivity analysis was performed to explore the impact of uncertainty on the results in terms of the assumptions, diagnostic characteristics and the clinical efficacy data used. The results were found to be very sensitive to sensitivity and specificity associated with formal case identification tools. Ideally probabilistic sensitivity analysis, which demonstrates the joint uncertainty between all of the different parameters used in the model, is also required. However, because of data limitations it was not possible to model the interaction between sensitivity and specificity associated with the Whooley questions or the PHQ-9; as a result probabilistic sensitivity analysis was not attempted.

One of the main limitations of the economic analysis is that, due to lack of available evidence, a number of the estimates used in the economic model were based on single studies and where necessary supplemented by the GDG expert opinion. For example, most validation data were for the EPDS strategy, and sensitivity and specificity for PHQ-9 and Whooley questions were based on single studies. Moreover, the available data for PHQ-9 that met the inclusion criteria were for antenatal period only. Nevertheless, this limitation was partially addressed by the extensive sensitivity analysis.

The utility weights incorporated in the analysis were for the general depression population and did not take into account the HRQoL of the infants, which is highly affected by their mothers' psychological mood. Also, the GDG felt that QALYs do not capture process characteristics associated with the interventions. The NICE *Guidelines Manual* (NICE, 2012a) recommends that non-direct health effect on individuals should be excluded in the NICE reference case and the perspective on outcomes should be all direct health effects. Nevertheless, the GDG felt that treatment interventions have an added value apart from the improvements in women's mental health and that these should be considered when making a recommendation.

The GDG also expressed a range of other concerns relating to the design of the analysis. For example, irrespective of the favourable findings associated with the

strategy utilising Whooley questions and PHQ-9 the GDG expressed their concern that a range of other mental health problems in women in the postnatal period would be missed since neither of the tools has been validated in identification of other mental health problems. The GDG also felt that Whooley questions and PHQ-9 should be part of a holistic approach to assess the mental health and the environment of the woman; it should act as a prompt and then clinical judgement should be used. The GDG also expressed their concern that recently the identification of women with depression in the perinatal period has decreased and that this is mainly due to women wishing to disguise information due to the fear of disclosing sensitive information. As a result, the GDG stressed the importance of building a trusting relationship, the attitude of staff, and the style of their approach when delivering case identification and the assessment review questions.

In summary, even though the use of Whooley questions followed by PHQ-9 was found to be the cost-effective approach in identifying depression in the postnatal period, the results were found to be sensitive to changes in diagnostic characteristics for formal case identification tools. This indicates that there is need for further research to compare the diagnostic performance of identification tools in women with depression in the postnatal period and in particular in women with other mental health problems in perinatal period; and also there is a need for more research relating to the pathways starting form identification and up to treatment.

Irrespective of the limitations, the findings of this model indicate the potential value associated with the systematic use of formal case identification tools in women with depression in the postnatal period.

Overall conclusions from the health economic evidence

Existing economic evidence is limited to identification methods for women with depression in the postnatal period. One existing UK-based study concluded that formal case identification was not cost-effective; however the study is characterised by potentially serious methodological limitations. International evidence is limited to one study conducted in New Zealand. The results suggested that a formal case identification programme is highly cost effective for depression in the postnatal period. Similarly, the economic analysis undertaken for this guideline suggests that for women with depression in the postnatal period the use of formal identification (such as, Whooley questions followed by PHQ-9) comprises a cost-effective strategy when compared with standard care case identification (GP assessment alone; without using formal identification tools) and also with strategies that do not utilise Whooley questions (use of EPDS only), because it appears to result in better outcomes (more women identified and higher number of QALYs) and lower total costs.

5.3.7 Linking evidence to recommendations

In developing recommendations for case identification, the GDG's primary concern was to ensure that women with a range of mental health problems in pregnancy and the postnatal period do not go unrecognised and therefore untreated. They were

concerned that, as highlighted in the review of experience of care in Chapter 6, some women may be unwilling to disclose or discuss any mental health problems because they are fearful that healthcare professionals might view them negatively in their role as a mother, or that their baby might be taken into care. Based on GDG expert clinical opinion there were also concerns that avoidance associated with some anxiety disorders, as well as drug or alcohol dependence, might also be a barrier to engagement.

In developing the recommendations the GDG had little data available on women in pregnancy and the postnatal period except for women who may have depression. As a consequence the GDG decided to use data on case identification in non-pregnant populations. The GDG considered this issue carefully and decided to draw on evidence from other NICE guidelines. However, there was sufficient evidence for depression to provide data on effectiveness of the various case identification tools and also to support development of the health economic model for case identification of depression. The model took into account the costs and consequences of not only correct identification but also the impact of false positives and false negatives. This meant that the model was able to inform aspects of the care pathway beyond initial case identification.

Based on evidence from non-pregnant populations, the GDG considered the possible benefits of a two-stage identification and diagnosis process: a brief and sensitive case identification tool suitable for use by a range of healthcare professionals in different settings; followed by a full clinical assessment, which may include a more detailed assessment instrument with better psychometric properties, for people with a positive response to the case identification questions. The benefit of using a brief case-finding approach in clinical settings where routine perinatal care takes place is not necessarily to diagnose depression or anxiety per se, but to reduce the number of women who need extensive assessment or evaluation with longer questionnaires such as the EPDS. Current NICE guidelines for depression (NICE, 2009a) recommend the use of the two Whooley questions. The questions do not require additional resources (such as copies of a questionnaire), and the value lies in part in their brevity and the fact that they lend themselves to use in both pregnancy and the postnatal period.

In supporting a recommendation for the use of case identification tools, the GDG considered the substantial costs associated with delayed diagnosis and management of unrecognised mental health problems in pregnancy and the postnatal period. The GDG recognised that early detection of mental health problems offers benefit to women who receive appropriate treatment for their condition, and may result in a considerable reduction in healthcare resource use and improvements in their HRQoL. Regarding depression in the postnatal period the guideline economic analysis suggested that the use of a brief case identification tool (that is, the Whooley questions), followed by the use of a more formal method (such as the EPDS or PHQ-9), appears to be the most cost-effective approach in the identification of depression in the postnatal period. The results were very sensitive to alternative scenarios

considered in the sensitivity analysis. The GDG took into account the fact that the results were determined based on very limited clinical data. Overall it seems that the strategies utilising a brief case identification tool (that is, the Whooley questions) are preferred to the strategies not utilising a brief case identification tool, however little can be said about which tool should be used for a more formal assessment (that is, the EPDS or PHQ-9). The GDG supported this model because its implications were broadly in line with recommendations made in other NICE guidelines for common mental health problems, and this would likely facilitate uptake of the recommendations. The additional depression identification question about the need for help (Arroll et al., 2005) was not included based on evidence that this third question had no conclusive benefit, and resulted in poor discrimination between true-negative and false-negative cases which may lead to an increased risk of depression being missed or lost to follow-up.

There was very limited diagnostic test accuracy data for the identification of anxiety disorders in pregnancy or the postnatal period or for specific anxiety disorders (including GAD, OCD, panic disorder, phobias, PTSD and social anxiety disorder) in pregnant or non-pregnant populations. However, the limited data available did not suggest that there were likely to be significant differences in the performance of these measures from that in the wider population on which previous NICE recommendations were based. For these reasons, the GDG judged that the use of the GAD-2 questions (and the additional use of the Generalized Anxiety Disorders scale – 7 items [GAD-7] or a question to elicit avoidance, if needed) was a reasonable extrapolation for pregnancy and the postnatal period. The GAD-2 has been found to have good diagnostic accuracy together with meeting clinical utility and feasibility criteria. As recommended in the current NICE guideline for common mental health disorders (NICE, 2011b), the GDG considered it important to add an additional question about avoidance to the two GAD questions in order to identify women with an established phobic disorder who might score low on the GAD-2 questions because they avoid phobic objects or situations and as a consequence of the avoidance would not experience significant anxiety or worry. In the absence of pregnancy or postnatal-specific validation and in the absence of any formal tools used for the identification of specific types of anxiety disorders, the use of the GAD-2 questions and additional question on avoidance was recommended to be considered as a tool to detect the range of anxiety disorders. However the GDG wished to draw attention to the range, prevalence and underdetection of anxiety disorders.

Based on both direct and extrapolated sensitivity and specificity data, the results of the health economic model for the identification of depression, and qualitative evidence and expert consensus opinion about the barriers to disclosure and engagement, the GDG recommended that a general discussion about mental health and wellbeing is had with all women upon first contact with primary care (or her booking visit) and during the early postnatal period, and that as part of this discussion the healthcare professional should consider asking the Whooley questions and the GAD-2. Positive response to either of the Whooley questions, a score of three or more on the GAD-2, or a positive response to the avoidance

question, should lead the healthcare professional to consider using the EPDS or the PHQ-9 in the case of depression, or the GAD-7 in the case of anxiety disorders, as part of a full assessment or referring the woman to her GP or to a mental health professional if a severe mental health problem is suspected. However, the GDG were concerned that the Whooley questions may still fail to identify depression for some women, therefore wished to recommend that even in the absence of a positive response to the depression identification questions, but where a woman is perceived to be at risk of a mental health problem or there is clinical concern, healthcare professionals should consider using a formal tool such as the EPDS or PHQ-9 as part of a full assessment.

The GDG recognised and wished to highlight that case identification should be an ongoing and individualised process during pregnancy and the postnatal period and that the most suitable healthcare professionals to perform this ongoing monitoring are those who have most contact with the woman, primarily health visitors, therefore a recommendation was made that at all subsequent contacts with a woman in pregnancy and the first postnatal year, the health visitor (and other healthcare professionals who have regular contact with the woman) should consider asking the Whooley questions and the GAD-2 as part of a general discussion about her mental health and wellbeing.

There was no high quality evidence for the case identification of severe mental illness in pregnancy and the postnatal period. However, the GDG wished to make recommendations in this area because of the need for healthcare professionals to act quickly in the event of postpartum psychosis. The GDG therefore agreed by consensus to recommend that at a woman's first contact with services, she should be asked about any past or present severe mental illness, previous or current treatment by a specialist mental health service and whether she has a first-degree relative with a history of severe perinatal mental illness. They also wished to urge healthcare professionals to be vigilant for possible symptoms of psychosis in women with any of these risk factors in the first two weeks after childbirth, and if a woman has sudden onset of symptoms suggesting postpartum psychosis, refer her to a secondary mental health service for immediate assessment (within 4 hours).

There was also no high quality evidence for the case identification of alcohol misuse in pregnancy and the postnatal period. The GDG wished to make a recommendation in this area given the risk of harm to the fetus, such as fetal alcohol syndrome. Therefore the GDG considered that the use of the Alcohol Use Disorders Identification Test (AUDIT), as specified in *Alcohol-Use Disorders* (NICE, 2011c), was suitable for use in pregnant women. For drug misuse in pregnant women, the GDG have cross-referred to the guideline on *Drug Misuse: Psychosocial Interventions* (NICE, 2007b).

The GDG reviewed recommendations from the previous 2007 guideline and judged that the advice on ensuring that information on any past or present mental health problem be shared with maternity services was still relevant. The recommendation

was reworded to conform to current NICE style and GDG expert judgement that it was important to specify the context and healthcare professional responsible for this so that all healthcare professionals referring a woman to a maternity service should ensure that communications with that service share information on any past and present mental health problem.

5.3.8 Recommendations

Recognising mental health problems in pregnancy and the postnatal period and referral

5.3.8.1 Recognise that women who have a mental health problem (or are worried that they might have) may be:

- unwilling to disclose or discuss their problem because of fear of stigma, negative perceptions of them as a mother or fear that their baby might be taken into care
- reluctant to engage, or have difficulty in engaging, in treatment because of avoidance associated with their mental health problem or dependence on alcohol or drugs. **[new 2014]**

5.3.8.2 All healthcare professionals referring a woman to a maternity service should ensure that communications with that service (including those relating to initial referral) share information on any past and present mental health problem. **[2014]**

Depression and anxiety disorders

5.3.8.3 Recognise that the range and prevalence of anxiety disorders (including generalised anxiety disorder, obsessive-compulsive disorder, panic disorder, phobias, post-traumatic stress disorder and social anxiety disorder) and depression are under-recognised throughout pregnancy and the postnatal period**. [new 2014]**

5.3.8.4 At a woman's first contact with primary care or her booking visit, and during the early postnatal period, consider asking the following depression identification questions as part of a general discussion about a woman's mental health and wellbeing:

- During the past month, have you often been bothered by feeling down, depressed or hopeless?
- During the past month, have you often been bothered by having little interest or pleasure in doing things?

Also consider asking about anxiety using the 2-item Generalized Anxiety Disorder scale (GAD-2):

- Over the last 2 weeks, how often have you been bothered by feeling nervous, anxious or on edge?[8]
- Over the last 2 weeks, how often have you been bothered by not being able to stop or control worrying?[9] **[new 2014]**

5.3.8.5 If a woman responds positively to either of the depression identification questions in recommendation 5.3.8.4 is at risk of developing a mental health problem, or there is clinical concern, consider:

- using the Edinburgh Postnatal Depression Scale (EPDS) or the Patient Health Questionnaire (PHQ-9) as part of a full assessment or
- referring the woman to her GP or, if a severe mental health problem is suspected, to a mental health professional. **[new 2014]**

5.3.8.6 If a woman scores 3 or more on the GAD-2 scale, consider:

- using the GAD-7 scale for further assessment or
- referring the woman to her GP or, if a severe mental health problem is suspected, to a mental health professional. **[new 2014]**

5.3.8.7 If a woman scores less than 3 on the GAD-2 scale, but you are still concerned she may have an anxiety disorder, ask the following question:

- Do you find yourself avoiding places or activities and does this cause you problems?

If she responds positively, consider:

- using the GAD-7 scale for further assessment or
- referring the woman to her GP or, if a severe mental health problem is suspected, to a mental health professional. **[new 2014]**

5.3.8.8 At all contacts after the first contact with primary care or the booking visit, the health visitor, and other healthcare professionals who have regular contact with a woman in pregnancy and the postnatal period (first year after birth), should consider:

- asking the 2 depression identification questions and the GAD-2 (see recommendation 5.3.8.4) as part of a general discussion about her mental health and wellbeing and
- using the EPDS or the PHQ-9 as part of monitoring. **[new 2014]**

Severe mental illness

5.3.8.9 At a woman's first contact with services in pregnancy and the postnatal period, ask about:

- any past or present severe mental illness
- past or present treatment by a specialist mental health service, including inpatient care
- any severe perinatal mental illness in a first-degree relative (mother, sister or daughter). **[2014]**

5.3.8.10 Refer to a secondary mental health service (preferably a specialist perinatal mental health service) for assessment and treatment, all women who:

- have or are suspected to have severe mental illness
- have any history of severe mental illness (during pregnancy or the postnatal period or at any other time).

Ensure that the woman's GP knows about the referral. **[new 2014]**

5.3.8.11 If a woman has any past or present severe mental illness or there is a family history of severe perinatal mental illness in a first-degree relative, be alert for possible symptoms of postpartum psychosis in the first 2 weeks after childbirth. **[new 2014]**

5.3.8.12 If a woman has sudden onset of symptoms suggesting postpartum psychosis, refer her to a secondary mental health service (preferably a specialist perinatal mental health service) for immediate assessment (within 4 hours of referral). **[new 2014]**

Alcohol and drug misuse

5.3.8.13 If alcohol misuse is suspected, use the Alcohol Use Disorders Identification Test (AUDIT) as an identification tool in line with recommendation 1.2.1.4 of the guideline on alcohol-use disorders (NICE clinical guideline 115). **[new 2014]**

5.3.8.14 If drug misuse is suspected, follow the recommendations on identification and assessment in section 1.2 of the guideline on drug misuse – psychosocial interventions (NICE clinical guideline 51). **[new 2014]**

5.3.9 Research recommendation

5.3.9.1 What methods can improve the identification of women at high risk of postpartum psychosis and reduce this risk?

[8] An answer of 'Not at all' scores 0; 'Several days' scores 1; 'More than half the days' scores 2; 'Nearly every day' scores 3.

[9] An answer of 'Not at all' scores 0; 'Several days' scores 1; 'More than half the days' scores 2; 'Nearly every day' scores 3.

5.4 ASSESSMENT

5.4.1 Introduction

Definition and aim of review

The review aims to identify the components and most appropriate structure of a diagnostic assessment for women with a mental health problem (any) in pregnancy and the postnatal period (defined in the this guideline as the first postnatal year).

5.4.2 Studies considered

The GDG was unable to identify any formal evaluations of the structure and content of the overall clinical assessment process for women with a possible mental health problem in pregnancy and the postnatal period other than the data on the various case identification instruments described above.

The GDG considered this topic to be important to the guideline, therefore they decided to draw on other sources of evidence to inform the development of recommendations in this area. These sources include:

- the reviews of the evidence and recommendations on assessment in the previous *Antenatal and Postnatal Mental Health* guideline (NICE, 2007a; NCCMH, 2007)
- the reviews of the evidence and recommendations on assessment in the existing NICE guidelines on specific mental health problems, including *Common Mental Health Disorders* (NICE, 2011b; NCCMH, 2011a) and *Psychosis and Schizophrenia* (NICE, 2014; NCCMH, 2014)
- reviews undertaken for this guideline, including case identification (see Section 5.3), experience of care (see Chapter 6) and pharmacological interventions (see Chapter 8)
- the expert knowledge and experience of the GDG.

5.4.3 Methodological approach

In drawing on the sources of evidence described above, the GDG was guided by the key principle that assessment and treatment of mental health problems in pregnancy and the postnatal period are not markedly different from assessment and treatment at other periods in a woman's life. However, there a number of factors specific to pregnancy and the postnatal period that requires the development of new recommendations or changes to existing recommendations, including: the health of the fetus or baby, the context in which the interventions are delivered, and specific variations in a woman's mental or physical health linked to pregnancy and the postnatal period. It follows from this principle that recommendations in the guideline should be made when evidence is identified and supports:

- a recommendation for an intervention that is unique to pregnancy or the postnatal period
- a recommendation to reflect the need for greater clarity about the use or application of interventions in an existing NICE guideline (including the previous *Antenatal and Postnatal Mental Health* guideline)
- a change to or modification of a recommendation for an intervention in an existing NICE guideline (including the previous *Antenatal and Postnatal Mental Health* guideline).

Having considered the clinical evidence and recommendations in other NICE guidelines, the experience of care review in chapter 6 of this guideline, and their own expert experience and opinion, the GDG then used informal consensus methods and the 'incorporate and adapt methodology' (as set out in Chapter 3) to determine recommendations.

5.4.4 Clinical evidence review (assessment)

When considering the reviews of the evidence and recommendations in other NICE guidelines, the GDG noted the commonality of the components for assessment for specific mental health problems, including common mental health problems such as depression and anxiety disorders, and severe mental illnesses such as psychosis and schizophrenia.

In order to provide a starting point for the development of recommendations, the GDG drew up a list of the following contextual and component factors of an assessment for women with a mental health problem in pregnancy and the postnatal period. This included:

- the stage of pregnancy (including the pre-conceptual period) and the postnatal period
- the needs of and concerns for the fetus or baby
- the setting in which the interventions are delivered and the need to ensure effective communication between all agencies involved in the assessment and care of the woman
- the need, where possible, to integrate case identification and assessment strategies
- the woman's symptom profile, including current and past symptoms, precipitating and maintaining factors, course and duration of current and past episodes, and family history
- social and personal functioning and current psychosocial stressors
- potential mental and physical comorbidities
- general physical health and side effects of medication
- potential involvement of a family member or carer to give a corroborative history
- treatment history and interventions that have been effective or ineffective in the past

- possible factors that may impact on the course of the mental health problem, including relationships, psychosocial factors and lifestyle changes
- social and economic issues that may be associated with the mental health problem
- risk to self and others
- the recognition that assessment is not a single time-limited intervention but is a continuing process throughout any period of care.

The GDG considered the factors set out above in light of both the evidence on case identification reviewed in Section 5.3 and recommendations in existing NICE guidelines. Based on this review the GDG concluded that new recommendations were needed for this guideline. Further evidence from the review of the experience of care (see Chapter 6) and reviews of the evidence on the efficacy of, and potentials harms associated with, interventions for mental health problems in pregnancy and the postnatal period (see Chapter 7 and 8), further informed the GDG in their development of recommendations for assessment.

In addition to the components and structure of the assessment, the GDG also discussed other processes and issues that would need to be considered around assessment or when planning treatment. These included:

- the need to take account of any learning disabilities or acquired cognitive impairments during assessment or subsequent treatment
- the need to develop a written care plan for a woman with a current or past severe mental illness
- the need for discussion with all women about any particular concerns they may have regarding the pregnancy and treatment for a mental health problem
- the need to seek specialist advice if the woman requests detailed discussion of risks and benefits of treatment
- the form that any discussion about likely risks and benefits of treatment should take, which should encompass acknowledging uncertainty about the magnitude of the risk of any specific intervention
- monitoring and increased contact, including for women who choose not to have, or stop, treatment for a mental health problem in pregnancy or the postnatal period
- the need for all healthcare professionals to understand the variations to the course and presentation of mental health problems in pregnancy and the postnatal period during assessment (and treatment).

5.4.5 Clinical evidence summary

The GDG was unable to identify any high-quality evidence that related to the process of assessment for women with a mental health problem in pregnancy and the postnatal period. As a result the GDG drew on the secondary sources of evidence

described in Section 5.4.2, their expert knowledge and experience and used informal consensus methods. The considerations that fed into the development of recommendations are described above and in Section 5.4.7.

5.4.6 Health economics evidence

No studies assessing the cost effectiveness of assessment systems for women with a mental health problem in pregnancy or the postnatal period were identified by the systematic search of the economic literature undertaken for this guideline. Details on the methods used for the systematic search of the economic literature are described in Chapter 3.

5.4.7 Linking evidence to recommendations

Relative value placed on the outcomes considered

When considering the development of the recommendations, the objective was to ensure that the specific contextual and clinical factors identified as important for women with a mental health problem in pregnancy and the postnatal period were taken into account so that an accurate assessment of a woman's needs and identification of the best available treatment or care option could be achieved.

Trade-off between benefits and harms

A central concern of the GDG was to ensure that the assessment adequately assessed the needs of the women and her fetus or baby, although the GDG also saw the value in making sure that the needs of her partner, family and carer were also adequately assessed. The focus in developing the recommendations was to address those areas where the evidence suggested that variations were needed to the usual care provided to the general population with a mental health problem. There is a risk that this could add to the burden of assessment and, in varying from routine practice, may be poorly implemented and lead to poorer outcomes. But the GDG judged that a number of factors such as the fear of disclosure of mental health problems in pregnancy (see Chapter 6), the concerns women have about the possible harms associated with the use of psychotropic medication in pregnancy, the risk of harm to the woman and fetus or baby of no or sub-optimal treatment, and the sudden and sometimes highly risky changes in mental state in pregnancy and the postnatal period, convinced the GDG of the need for specific recommendations in the area of assessment. The recommendation on what an assessment for a woman with a mental health problem in pregnancy and the postnatal period should cover was based on the discussion of the evidence outlined in Section 5.4.4. In addition, stakeholder comments highlighted that a woman's attitude towards pregnancy (including denial of pregnancy), the woman's experience of pregnancy and any problems experienced by her or the fetus or baby, and the mother-baby relationship were important components of an assessment and diagnosis of a suspected mental health problem in pregnancy and the postnatal period. As stated in Section 5.4.4, the GDG saw many commonalities in the assessment of mental health problems in other NICE guidelines and did not see the value of making separate recommendations for different mental health problems. Having said that, the GDG took account of the fact that most

women are first seen (and many effectively treated) in non-specialist mental health settings. The GDG therefore decided to structure the assessment recommendations in a way that reflected this. The GDG also saw the value in highlighting that all healthcare professionals should understand the variations in the presentation and course of mental health problems in pregnancy and the postnatal period and the context in which they are often treated (for example, maternity services). In addition, one recommendation from *Common Mental Health Disorders* on a stepped care model of delivery was judged by the GDG to be relevant to the delivery of interventions in this guideline on antenatal and postnatal mental health. Therefore the GDG recommended the use of stepped care and cross-referred to the *Common Mental Health Disorders* guideline for further information.

In addition the GDG wished to make specific recommendations to urge healthcare professionals to take account of learning disabilities or acquired cognitive impairments when assessing (or treating) a mental health problem in pregnancy or the postnatal period. The GDG was also aware of the potential risks for the fetus or baby that might arise from the mother's mental health problem and the fact that this would require not only careful assessment of risk but also effective communication with a range of agencies. The GDG judged that women with a current or past severe mental illness should have a written care plan in place.

The GDG was aware that assessment and the monitoring of the effects of interventions should be a continual process and as far as possible integrated into routine care. This should start with a more detailed assessment following initial identification but should also support more detailed disorder-specific monitoring of mental state.

For any woman with a mental health problem, whether it is pre-existing or has developed in pregnancy or the postnatal period, discussion about treatment or prevention options in pregnancy and the postnatal period need to cover the likely benefits and harms associated with treatment, and what might happen if the woman decides not to have treatment or she stops or changes psychotropic medication abruptly. In developing these recommendations the GDG was also mindful that some of the recommendations required specialist knowledge (for example, of the trade-off of harms and benefits associated with the use of psychotropic medication). Recommendations to seek specialist advice were therefore made, which also detail the form that the discussion should take, which should acknowledge the uncertainty about the magnitude of the risk of any specific intervention. The GDG was keen to support the active involvement of the women in all decisions about her care (including in the pre-conceptual phase) and encompassed this in the recommendations.

Trade-off between net health benefits and resource use

No studies assessing the cost effectiveness of assessment systems for women with a mental health problem in pregnancy or the postnatal period were identified, however the GDG acknowledged that appropriate assessment enables women to

receive suitable treatment according to their needs, thus ensuring efficient use of available healthcare resources. The GDG also considered the cost of providing such assessment to be small (for example, the cost of health visitor consultation ranges from £49 to £71 per hour) relative to the substantial costs associated with delayed diagnosis and management of unrecognised and/or misdiagnosed mental health problems in pregnancy or the postnatal period, no or sub-optimal treatment, and the potential risks for the fetus or baby that might arise from under-recognition of mother's mental health problem.

Quality of the evidence

No high-quality evidence was identified that examined the structure and content of the overall clinical assessment process for women in pregnancy and the postnatal period. The recommendations were therefore based on a review of existing NICE guidelines, reviews undertaken for this guideline and the expert opinion of the GDG.

5.4.8 Recommendations

Using this guideline in conjunction with other NICE guidelines

Assessment and treatment in pregnancy and the postnatal period

5.4.8.1 Use this guideline in conjunction with the NICE guideline for a specific mental health problem (see the related NICE guidance in section 3.2 [in the NICE guideline]) to inform assessment and treatment decisions in pregnancy and the postnatal period, and take into account:

- any variations in the nature and presentation of the mental health problem in pregnancy or the postnatal period
- the setting for assessment and treatment (for example, primary or secondary care services or in the community, the home or remotely by phone or computer)
- recommendations 5.4.8.5 5.4.8.5 to5.4.8.10 in this guideline on assessment in pregnancy and the postnatal period
- recommendations 8.9.1.6 to 8.9.1.33 in this guideline on starting, using and stopping treatment in pregnancy and the postnatal period
- recommendations 5.4.8.13 to 5.4.8.14, 7.7.1.6 to 7.7.1.17 and 8.9.1.35 to 8.9.1.48 in this guideline on treating specific mental health problems in pregnancy and the postnatal period. **[new 2014]**

Principles of care in pregnancy and the postnatal period

5.4.8.2 **Supporting women and their partners, families and carers**Take into account and, if appropriate, assess and address the needs of partners, families and carers that might affect a woman with a mental health problem in pregnancy and the postnatal period. These include:

- the welfare of the baby and other dependent children and adults
- the role of the partner, family or carer in providing support
- the potential effect of any mental health problem on the woman's relationship with her partner, family or carer. **[new 2014]**

Treatment decisions, advice and monitoring for women who are planning a pregnancy, pregnant or in the postnatal period

Monitoring and increased contact

5.4.8.3 Healthcare professionals working in universal services and those caring for women in mental health services should:

- assess the level of contact and support needed by women with a mental health problem (current or past) and those at risk of developing one

- agree the level of contact and support with each woman, including those who are not having treatment for a mental health problem
- monitor regularly for symptoms throughout pregnancy and the postnatal period, particularly in the first few weeks after childbirth. **[new 2014]**

5.4.8.4 Discuss and plan how symptoms will be monitored (for example, by using validated self-report questionnaires, such as the Edinburgh Postnatal Depression Scale [EPDS], Patient Health Questionnaire [PHQ-9] or the 7-item Generalized Anxiety Disorder scale [GAD-7]). **[new 2014]**

Assessment and care planning in pregnancy and the postnatal period

5.4.8.5 Assessment and diagnosis of a suspected mental health problem in pregnancy and the postnatal period should include:

- history of any mental health problem, including in pregnancy or the postnatal period
- physical wellbeing (including weight, smoking, nutrition and activity level) and history of any physical health problem
- alcohol and drug misuse
- the woman's attitude towards the pregnancy, including denial of pregnancy
- the woman's experience of pregnancy and any problems experienced by her, the fetus or the baby
- the mother–baby relationship
- any past or present treatment for a mental health problem, and response to any treatment
- social networks and quality of interpersonal relationships
- living conditions and social isolation
- family history (first-degree relative) of mental health problems
- domestic violence and abuse, sexual abuse, trauma or childhood maltreatment
- housing, employment, economic and immigration status
- responsibilities as a carer for other children and young people or other adults. **[new 2014]**

5.4.8.6 When assessing or treating a mental health problem in pregnancy or the postnatal period, take account of any learning disabilities or acquired cognitive impairments, and assess the need to consult with a specialist when developing care plans. **[new 2014]**

5.4.8.7 Carry out a risk assessment in conjunction with the woman and, if she agrees, her partner, family or carer. Focus on areas that are likely to present possible risk such as self-neglect, self-harm, suicidal thoughts and intent,

risks to others (including the baby), smoking, drug or alcohol misuse and domestic violence and abuse. **[new 2014]**

5.4.8.8 If there is a risk of, or there are concerns about, suspected child maltreatment, follow local safeguarding protocols. **[new 2014]**

5.4.8.9 If there is a risk of self-harm or suicide:

- assess whether the woman has adequate social support and is aware of sources of help
- arrange help appropriate to the level of risk
- inform all relevant healthcare professionals (including the GP and those identified in the care plan) [see recommendation 5.4.8.10]
- advise the woman, and her partner, family or carer, to seek further help if the situation deteriorates. **[new 2014]**

5.4.8.10 Professionals in secondary mental health services, including specialist perinatal mental health services, should develop a written care plan in collaboration with a woman who has or has had a severe mental illness. If she agrees, her partner, family or carer should also be involved. The plan should cover pregnancy, childbirth and the postnatal period (including the potential impact of the illness on the baby) and should include:

- a clear statement of jointly agreed treatment goals and how outcomes will be routinely monitored
- increased contact with and referral to specialist perinatal mental health services
- the names and contact details of key professionals.

The care plan should be recorded in all versions of the woman's notes (her own records and maternity, primary care and mental health notes) and a copy given to the woman and all involved professionals. **[new 2014]**

Providing interventions in pregnancy and the postnatal period

5.4.8.11 All healthcare professionals providing assessment and interventions for mental health problems in pregnancy and the postnatal period should understand the variations in their presentation and course at these times, how these variations affect treatment, and the context in which they are assessed and treated (for example, maternity services, health visiting and mental health services). **[new 2014]**

5.4.8.12 Provide interventions for mental health problems in pregnancy and the postnatal period within a stepped-care model of service delivery in line with

recommendation 1.5.1.3 of the guideline on common mental health disorders (NICE clinical guideline 123). **[new 2014]**

Treating specific mental health problems in pregnancy and the postnatal period

Interventions for alcohol and drug misuse

5.4.8.13 If hazardous drug or alcohol misuse is identified in pregnancy or the postnatal period, refer or offer brief interventions in line with section 1.3.1 of the guideline on drug misuse – psychosocial interventions (NICE clinical guideline 51) or the guideline on alcohol-use disorders: preventing harmful drinking (NICE public health guidance 24). **[new 2014]**

5.4.8.14 If harmful or dependent drug or alcohol misuse is identified in pregnancy or the postnatal period, refer the woman to a specialist substance misuse service for advice and treatment. **[new 2014]**

6 EXPERIENCE OF CARE

6.1 INTRODUCTION

The focus of this chapter is the experience of care of women who have an existing mental health problem or who develop one in pregnancy or the postnatal period (from childbirth up to 1 year), although it is potentially relevant to all women and girls of childbearing potential (because any could in principle develop a mental health problem). A thematic analysis of the qualitative literature was undertaken in order to identify themes relevant to the experience of care for women with a mental health problem in pregnancy or the postnatal period. This analysis directly informs the development of recommendations in this chapter aiming to improve women's experience of care, and the experience of their partners, families and carers, but it also informs the development of other recommendations in the guideline.

Many aspects of treatment and the principles underpinning good care are common to all people in receipt of healthcare, including women with a mental health problem in pregnancy or the postnatal period. Relevant NICE guidance sets out the principles for improving the experience of care for people using adult NHS mental health services (*Service User Experience in Adult Mental Health* [NICE, 2011d; NCCMH, 2012]) and general medical services (*Patient Experience in Adult NHS Services* [NICE, 2012b; National Clinical Guideline Centre, 2012]). *Service User Experience in Adult Mental Health* guidance examined the evidence for improving experience of mental health services in seven main areas: access to community care, assessment (non-acute), community care, assessment and referral in crisis, hospital care, discharge and transfer of care and detention under the Mental Health Act. The *Patient Experience in Adult NHS Services* guidance examined the evidence for improving experience of adult health services in five main areas: the patient as an individual, the essential requirements of care, the tailoring of healthcare services for each patient, continuity of care and relationships and enabling patients to actively participate in their care.

However, there are a number of factors (described in detail in the introduction), including the impact on the fetus or baby of the mother's mental health and use of psychotropic medication, that are unique to pregnancy and the postnatal period and that alter women's experience of healthcare. At other times, when the woman is not pregnant or caring for her baby, the sole focus of care and treatment is the woman, but in pregnancy and the postnatal period, the emphasis shifts to a concern for the fetus and baby as well as the woman which can contribute to different and difficult experiences of care particularly where the needs of the mother and fetus or baby conflict.

Therefore while it is expected that health and social care professionals will consult *Service User Experience in Adult Mental Health* and *Patient Experience in Adult NHS*

Services to improve all aspects of experience across the care pathway for adults using mental health services, there are specific areas of concern for women with a mental health problem in pregnancy and the postnatal period that need to be addressed by the current guideline.

The large majority of women with a mental health problem in pregnancy and the postnatal period will be identified and treated in primary care with no or only limited input or advice from specialist mental health services. Another group of women will not have their problem recognised at all and so will not access treatment. This lack of recognition stems from a number of factors including a historical focus on mental health problems in the postnatal period as opposed to in pregnancy and a concern on the part of some women about disclosing any mental health problem particularly due to fears about loss of custody. Understanding women's experience of recognition of their mental health problem and the context in which it is undertaken is a vital first step in providing effective treatment.

A mother's concerns about the possible impact of a mental health problem on the fetus or baby and the benefits or possible harms associated with treatment, may outweigh her concerns for her own health. A better understanding of these concerns and about how they may be sensitively addressed is also important when establishing effective treatment plans.

Those women who develop a severe mental illness in pregnancy or the postnatal period require treatment in a secondary mental health service or specialist perinatal mental health service. It is important that their experience is also captured to improve potential areas of concern, such as how all of the services and agencies involved (for example, primary, maternity and mental health and social care) can communicate and work effectively with each other.

Current practice

There is currently considerable variation in the experience of women with a mental health problem in pregnancy and the postnatal period. This may arise from the concerns outlined above but may also relate to other factors including: limited staff training or knowledge; the absence of tools or systems to support the recognition of mental health problems and ensure effective communication; and the limited availability of specialist services to provide advice or treatment for more severely ill women. As a result many women may go to voluntary sector organisations such as 'Netmums' for information and support. While such organisations play a vital role in enabling women to access informal support, not all women access them and their existence does not remove the responsibility for health services to ensure that the care of women with mental health problems in pregnancy and the postnatal period is a positive experience with access to and engagement with the best available treatment.

6.2 REVIEW OF THE PRIMARY EVIDENCE

6.2.1 Clinical review protocol (experience of care)

The review protocol, including the review questions, information about the databases searched, and the eligibility criteria used for this section of the guideline, can be found in Table 25 (further information about the search strategy can be found in Appendix 10). A systematic search for published reviews of relevant qualitative studies of women with mental health problems in pregnancy or the postnatal period was undertaken using standard NCCMH procedures as described in Chapter 3. Reviews were sought of qualitative studies that used relevant first-hand experiences. The GDG did not specify a particular outcome. Instead the review was concerned with any narrative data that highlighted the experience of care. Where a significant body of systematic reviews was not identified, the GDG looked for primary studies and adopted the method described in Chapter 3, Section 3.5.2, for the analysis of the studies.

Table 25: Databases searched and inclusion/exclusion criteria for clinical evidence

Component	Description
Review question (s)	What factors prevent women with a mental health problem who are pregnant or in the postnatal period accessing mental healthcare services? (RQ1.1) What factors improve or diminish the experience of services for women with a mental health problem who are pregnant or in the postnatal period? (RQ1.2) What modifications to services improve the experience of using services for women with a mental health problem who are pregnant or in the postnatal period? (RQ1.3)
Sub-question (s)	For women with mental health problems who are pregnant or in the postnatal period, is the experience of care different for: black and minority ethnic groups socioeconomic groups asylum seekers and refugees women who are victims of trafficking women with learning and physical disabilities gypsies and travellers women in prison?
Objectives	To identify obstacles to access by synthesising qualitative evidence and through expert consensus. To identify factors that improve or diminish the experiences of health and social services for women with a mental health problem in pregnancy or the postnatal period. To evaluate the effectiveness of interventions for improving the experience of health and social services for women with a mental health problem in pregnancy or in the postnatal period.
Criteria for considering studies for the review	
Population	Included

	Women who are pregnant and in the postnatal period (from childbirth up to 1 year): with subthreshold symptoms of a mental health problem who are 'at risk' of developing a mental health problem with existing mild, moderate and severe mental health problems who are currently receiving treatment (psychological or pharmacological) for an existing mental health problem Excluded women with a mental health problem after the first postnatal year women who are not pregnant or in the postnatal period (from childbirth up to 1 year) If some, but not all, of a study's participants are eligible for review, the study authors will be contacted for disaggregated data. If appropriate disaggregated data cannot be obtained, then a study will be included if the majority (at least 51%) of its participants are eligible for the guideline review. Women who are more than 1 year into the postnatal period but are giving retrospective reports of the immediate postnatal period (within 1 year after childbirth) will also be included.
Intervention	Review question 1.1 Factors or attributes of the individual who requires mental healthcare, that can inhibit access to services Practitioner-level factors or attributes that can inhibit an individual from accessing healthcare Excluded factors Systems and processes Practical or resource-based factors Review question 1.2 Actions by services that could improve or diminish the experience of care for example: Form, frequency and content of interactions with service users, families, carers or peers Sharing information with and receiving information from service users, families, carers or peers Planning of care with service users, families, carers or peers Review question 1.3 Any intervention delivered directly to the service user, families, carers or peers. The provision of financial and practical support (for example direct payments) is outside of the scope of this guideline and will not be included. This review will exclude: experiences of mental health problems in pregnancy or the postnatal period with no explicit implications for management, planning and/or delivery of care; case studies; autobiographical accounts; and qualitative measures of perceived intervention effectiveness where a quantitative approach would have been more appropriate.
Comparison	None

Critical outcomes	Review question 1.1 Identified factors affecting access Review question 1.2 Themes and specific issues that service users identify as improving or diminishing their experience of healthcare services Review question 1.3 Service user: Engagement, acceptability and uptake of services Retention Quality of Life Satisfaction (validated measures only, specific items will not be analysed).
Time points	Not applicable.
Study design	Review question 1.1 and 1.2 Systematic reviews of qualitative studies, primary qualitative studies, surveys. Review question 1.3 RCTs Systematic reviews of RCTs Systematic reviews of qualitative studies, primary qualitative studies, surveys. Books, dissertation abstracts, trade magazines, policy and guidance, non-English language papers, and non-empirical research will be excluded.
Include unpublished data?	Yes but only where: the evidence was accompanied by a report containing sufficient detail to properly assess the quality of the data the evidence was submitted with the understanding that data from the study and a summary of the study's characteristics will be published in the full guideline. Therefore, the GDG should not accept evidence submitted as commercial in confidence. However, the GDG should recognise that unpublished evidence submitted by investigators, might later be retracted by those investigators if the inclusion of such data would jeopardise publication of their research.
Restriction by date?	Systematic reviews of qualitative studies, primary qualitative studies, surveys: 1995 to 7 April 2014 (new search for this guideline) Systematic reviews of RCTs, RCTs: 2006 to 7 April 2014 (update search post-2007 guideline)
Minimum sample size	Include all sample sizes greater than one
Study setting	UK primary, secondary and tertiary healthcare services relevant to the NHS. This guideline will also be relevant to the work of, but will not provide specific recommendations to, NHS funded services (for example, social services, or the non-statutory sector).
Search strategy	Review question: 1.1, 1.2 ,1.3 Study design searched: Systematic reviews of qualitative studies, primary qualitative studies, surveys. Databases searched: General medical databases: CINAHL, Embase, MEDLINE, PreMEDLINE, PsycINFO

	Date restrictions: 1995 to 7 April 2014 **Review question: 1.3** Study designs searched: RCTs, systematic reviews of RCTs Databases searched: General medical databases: CINAHL, Embase, MEDLINE, PreMEDLINE, PsycINFO Topic specific databases: CDSR, CENTRAL, DARE, HTA Date restrictions: 2006 to 7 April 2014
Searching other resources	Hand-reference searching of retrieved literature
Review strategy	Review question 1.1 and 1.2 Thematic synthesis of qualitative papers. A modified matrix of service user experience will be used to organise themes. Review question 1.3 The initial aim is to conduct a meta-analysis evaluating the clinical effectiveness of the interventions. High quality systematic reviews (for example, Cochrane reviews) identified as part of the search will be utilised but will only be used if they meet the following criteria: methodology of the review is deemed appropriate and is in keeping with guideline methods PICO of the review is relevant to the guideline the review is of a high quality without substantial errors that could have an impact on conclusions and guideline recommendations. For each review, the following will also be extracted: year of review; total number of study participants; inclusion and exclusion criteria; age (mean); race (percent white); diagnosis. For each intervention or comparison group of interest, dose, frequency and duration of interventions will also be extracted.

6.2.2 Introduction

A search for systematic reviews of the experience of care of women with a mental health problem in pregnancy and the postnatal period was conducted. However, no relevant systematic reviews were considered suitable for inclusion. Consequently, a second search was conducted to identify relevant primary qualitative studies and survey data. The literature review supported a thematic analysis of the qualitative data reported in the primary studies.

6.2.3 Method

The method used in this section is set out in Chapter 3. In summary, the included primary qualitative studies (see Table 25 for details of inclusion criteria) were reviewed using data extraction techniques consistent with the methodology used in *Service User Experience in Adult Mental Health* (NICE, 2011d; NCCMH, 2012). Each included study was reviewed by members of the review team and broad themes were identified and coded using the matrix detailed in Service User Experience in

Adult Mental Health. This matrix was formed by creating a table with the eight dimensions of person-centred care developed by the Picker Institute Europe10, down the vertical axis, and the key points on a pathway of care (as specified by the GDG) across the horizontal axis (see Table 27). The Picker Institute's dimensions of patient-centred care were chosen because they are well established, comprehensive, and based on research. In addition, a variation of these dimensions has been adopted by the US Institute of Medicine (Institute of Medicine, 2001). Consultation with another reviewer or members of the GDG was used to overcome difficulties with coding. Data from studies was extracted independently by two reviewers. Disagreements were resolved through discussion. Where consensus could not be reached, a third reviewer or GDG member resolved the disagreement. Masked assessment (that is, blind to the journal from which the article comes, the authors, the institution and the magnitude of the effect) was not used since it is unclear that doing so reduces bias (Jadad et al., 1996; Berlin, 2001). The superordinate and subordinate themes identified through the thematic synthesis of primary qualitative papers are used as headings and sub-headings to organise the evidence review below (Section 6.2.5).

6.2.4 Qualitative studies considered

In the search, 189 studies met the eligibility criteria for full-text retrieval. Of these, 39 provided relevant clinical evidence and were included in the review: ANTONYSAMY2009 (Antonysamy et al., 2009), AYERS2006 (Ayers et al., 2006), BOATH2004 (Boath et al., 2004), BREUSTEDT2013 (Breustedt & Puckering, 2013), CHEWGRAHAM2009 (Chew-Graham et al., 2009), COOKE2012 (Cooke et al., 2012), DEJONGE2001 (de Jonge, 2001), EDGE2005/2007/2008 (one study reported across three papers: Edge & Rogers, 2005; Edge, 2007; Edge, 2008), EDGE2011 (Edge, 2011), EDWARDS2005 (Edwards & Timmons, 2005), HALL2006 (Hall, 2006), HANLEY2006 (Hanley & Long, 2006), HERON2012 (Heron et al., 2012), HUNT2009 (Hunt et al., 2009), MAPP2005A/2005B (Mapp & Hudson, 2005a; Mapp, 2005b), MCCREIGHT2008 (McCreight, 2008), MCGRATH2013 (McGrath et al., 2013), NICHOLLS2007 (Nicholls & Ayers, 2007), PARVIN2004 (Parvin et al., 2004), PATEL2013 (Patel et al., 2013), RAYMOND2009 (Raymond, 2009), ROBERTSON2003 (Robertson & Lyons, 2003), RYNINKS2014 (Ryninks et al., 2014), SHAKESPEARE2003 (Shakespeare et al., 2003), SHAKESPEARE2006 (Shakespeare et al., 2006), SIMMONS2006 (Simmons et al., 2006), SLADE2010 (Slade et al., 2010), SMITH2007 (Smith & Gibb, 2007), SNOWDON2012 (Snowdon et al., 2012), STANLEY2006 (Stanley et al., 2006), STAPLETON2008 (Stapleton et al., 2008), TEMPLETON2003 (Templeton et al., 2003), THOMSON2008 (Thomson & Downe, 2008), THOMSON2013 (Thomson & Downe, 2013), THURTLE2003 (Thurtle, 2003), TSARTSARA2002 (Tsartsara & Johnson, 2002), TURNER2008 (Turner et al., 2008), TURNER2010 (Turner et al., 2010), WITTKOWSKI2011 (Wittkowski et al., 2011). All studies were published in peer-reviewed journals between 2001 and 2014.

[10] http://www.pickereurope.org/patientcentred

One hundred and fifty studies were excluded from the analysis. The most common reasons for exclusion were: non-UK setting for the study; the paper was a systematic review with no new useable data; the paper was concerned with the experience of the mental health problem itself with no explicit implications for management, planning and/or delivery of care; or the outcomes were not mental health focused. Further information about both included and excluded studies can be found in Appendix 18.

The characteristics of the included primary qualitative studies have been summarised in Table 26, the quality of these studies is summarised in Table 27 and Table 28 and the studies from which data were extracted are summarised in the experience of care matrix in Table 29, categorised according to the key themes

Table 26: Study information table for included primary qualitative studies of the experience of care for women with a mental health problem in pregnancy or the postnatal period

	Primary qualitative studies of the experience of care of women with a mental health problem in pregnancy or the postnatal period
Included studies	K=39
Sample size	4-280 (mean: 24)
Age of women (years)	17-60 (mean: 32) [includes retrospective account of experiences]
Age of child (months)	0.5-280 (mean: 26) [includes retrospective account of experiences]
Ethnicity (% white)	0-100 (mean: 67.5)
Diagnosis	Postnatal depression (K=13; 33%); antenatal depression (K=1; 3%); postnatal and/or antenatal depression (K=2; 5%); postpartum psychosis (K=4; 10%); PTSD (K=2; 5%); multiple (K=2; 5%); eating disorder (K=1; 3%); substance misuse (K=1; 3%)
Primiparous (%)	33-100 (mean: 59.5)
Method of delivery (%)	Vaginal (natural): 17-89 (mean: 52.1); vaginal (assisted): 5-28 (mean: 14.3); caesarean: 11-100 (mean: 38.7)
Focus of study	Barriers to access (K=12; 31%); factors that diminish the experience of care (K=5; 13%); experience of traumatic birth/obstetric emergency (K=4; 10%); factors that improve the experience of care (K=3; 8%); experience of antidepressants (K=3; 8%); experience of an inpatient unit (K=2; 5%); experience of listening visits (K=2; 5%); experience of post-miscarriage information and support (K=2; 5%); experience of routine screening with the EPDS (K=1; 3%); experience of specialist health visiting service (K=1; 3%); experience of termination of pregnancy following diagnosis of fetal abnormality (K=1; 3%); experience of stillbirth (K=1; 3%); experience of pregnancy loss due to miscarriage or stillbirth (K=1; 3%); modifications that improve the experience of care (K=1; 3%)
Data collection method	Face-to-face interview (K=25; 64%); interview (format not reported; K=8; 21%); focus group (K=3; 8%); questionnaire (open-ended) (K=2; 5%); focus group and interview (K=1; 3%)
Setting	Home (K=20; 51%); not reported (K=12; 31%); multiple (home, community settings, hospital; K=4; 10%); community setting (K=2; 5%); postal questionnaire (K=1; 3%)

Table 27: Quality of included studies for service user experience (part 1)

Study ID	Key research question/aim	Theoretical approach		Study design	Data collection	Validity	
		Is a qualitative approach appropriate?	Is the study clear in what it seeks to do?	Defensible/ rigorous methodology?	How well was the data collection carried out?	Is the context clearly described?	Were the methods reliable?
ANTONYSAMY2009	Experience of inpatient unit	Appropriate	Clear	Defensible	Appropriate	Clear	Reliable
AYERS2006	Factors that diminish EoC	Appropriate	Clear	Defensible	Appropriate	Clear	Not sure[1]
BOATH2004	Experience of antidepressants	Appropriate	Clear	Defensible	Appropriate	Clear	Not sure[1]
BREUSTEDT2013	Factors that improve EoC	Appropriate	Clear	Defensible	Appropriate	Clear	Not sure[1]
CHEWGRAHAM2009	Barriers to access	Appropriate	Clear	Defensible	Appropriate	Unclear[2]	Not sure[1]
COOKE2012	Barriers to access	Appropriate	Clear	Defensible	Appropriate	Unclear[3]	Not sure[1]
DEJONGE2001	Barriers to access	Appropriate	Clear	Defensible	Appropriate	Clear	Reliable
EDGE2005/2007/2008	Barriers to access	Appropriate	Clear	Defensible	Appropriate	Clear	Not sure[1]
EDGE2011	Barriers to access	Appropriate	Clear	Defensible	Appropriate	Clear	Not sure[1]
EDWARDS2005	Barriers to access	Appropriate	Clear	Defensible	Appropriate	Unclear[2]	Not sure[1]
HALL2006	Barriers to access	Appropriate	Clear	Defensible	Appropriate	Clear	Not sure[1]
HANLEY2006	Factors that improve EoC	Appropriate	Clear	Defensible	Appropriate	Clear	Not sure[1]
HERON2012	Experience of inpatient unit	Appropriate	Clear	Defensible	Appropriate	Unclear[3]	Not sure[1]

HUNT2009	Experience of termination of pregnancy following diagnosis of fetal abnormality	Appropriate	Clear	Defensible	Appropriate	Unclear[2]	Not sure[1]
MAPP2005A/2005B	Experience of obstetric emergency	Appropriate	Clear	Defensible	Appropriate	Unclear[2]	Not sure[1]
MCCREIGHT2008	Experience of pregnancy loss due to stillbirth or miscarriage	Appropriate	Clear	Defensible	Appropriate	Clear	Reliable
MCGRATH2013	Factors that diminish EoC	Appropriate	Clear	Defensible	Appropriate	Clear	Not sure[1]
NICHOLLS2007	Factors that diminish EoC	Appropriate	Clear	Defensible	Appropriate	Clear	Not sure[1]
PARVIN2004	Barriers to access	Appropriate	Clear	Defensible	Appropriate	Unclear[2]	Not sure[1]
PATEL2013	Experience of antidepressants	Appropriate	Clear	Defensible	Appropriate	Clear	Not sure[1]
RAYMOND2009	Modifications that improve EoC	Appropriate	Clear	Defensible	Appropriate	Clear	Not sure[1]
ROBERTSON2003	Factors that diminish EoC	Appropriate	Clear	Defensible	Appropriate	Clear	Not sure[1]
RYNINKS2014	Experience of stillbirth	Appropriate	Clear	Defensible	Appropriate	Clear	Not sure[1]
SHAKESPEARE2003	Experience of routine screening with EPDS	Appropriate	Clear	Defensible	Appropriate	Clear	Not sure[1]
SHAKESPEARE2006	Experience of listening visits	Appropriate	Clear	Defensible	Appropriate	Clear	Not sure[1]
SIMMONS2006	Experience of post-miscarriage information and support	Appropriate	Clear	Defensible	Appropriate	Unclear[2]	Not sure[1]

SLADE2010	Factors that improve EoC	Appropriate	Clear	Defensible	Appropriate	Unclear[3]	Not sure[1]
SMITH2007	Experience of a specialist health visiting service	Appropriate	Clear	Defensible	Appropriate	Clear	Not sure[1]
SNOWDON2012	Experience of traumatic birth	Appropriate	Clear	Defensible	Appropriate	Unclear[2,3]	Not sure[1]
STANLEY2006	Barriers to access	Appropriate	Clear	Defensible	Appropriate	Unclear[2]	Not sure[1]
STAPLETON2008	Factors that diminish EoC	Appropriate	Clear	Defensible	Appropriate	Clear	Not sure[1]
TEMPLETON2003	Barriers to access	Appropriate	Clear	Defensible	Appropriate	Unclear[2]	Reliable
THOMSON2008	Experience of traumatic birth	Appropriate	Clear	Defensible	Appropriate	Clear	Not sure[1]
THOMSON2013	Experience of traumatic birth	Appropriate	Clear	Defensible	Appropriate	Clear	Not sure[1]
THURTLE2003	Barriers to access	Appropriate	Clear	Defensible	Appropriate	Clear	Not sure[1]
TSARTSARA2002	Experience of post-miscarriage information and support	Appropriate	Clear	Defensible	Appropriate	Unclear[2]	Not sure[1]
TURNER2008	Experience of antidepressants	Appropriate	Clear	Defensible	Appropriate	Clear	Not sure[1]
TURNER2010	Experience of listening visits	Appropriate	Clear	Defensible	Appropriate	Clear	Not sure[1]
WITTKOWSKI2011	Barriers to access	Appropriate	Clear	Defensible	Appropriate	Clear	Not sure[1]

Notes. [1] Data were collected with only one method
[2] Description of participant characteristics is very limited
[3] Setting not reported

Table 28: Quality of included studies for service user experience (part 2)

Study ID	Analysis				Ethics	
	Are the data 'rich'?	Is the analysis reliable?	Are the findings convincing?	Are the conclusions adequate?	Was the study approved by an ethics committee?	Is the role of the researcher clearly described?
ANTONYSAMY2009	Rich	Not sure/not reported[1]	Convincing	Adequate	Not sure/not reported/not applicable[2]	Not sure/not reported[3]
AYERS2006	Rich	Reliable	Convincing	Adequate	Yes	Not sure/not reported[3]
BOATH2004	Rich	Not sure/not reported[1]	Convincing	Adequate	Yes	Not sure/not reported[3]
BREUSTEDT2013	Rich	Not sure/not reported[1]	Convincing	Adequate	Yes	Not sure/not reported[3]
CHEWGRAHAM2009	Rich	Reliable	Convincing	Adequate	Yes	Not sure/not reported[3]
COOKE2012	Rich	Reliable	Convincing	Adequate	Yes	Not sure/not reported[3]
DEJONGE2001	Rich	Not sure/not reported[1]	Convincing	Adequate	Not sure/not reported/not applicable[2]	Not sure/not reported[3]
EDGE2005/2007/2008	Rich	Not sure/not reported[1]	Convincing	Adequate	Yes	Not sure/not reported[3]
EDGE2011	Rich	Not sure/not reported[1]	Convincing	Adequate	Yes	Not sure/not reported[3]
EDWARDS2005	Rich	Not sure/not reported[1]	Convincing	Adequate	Yes	Clear

HALL2006	Rich	Reliable	Convincing	Adequate	Yes	Not sure/not reported[3]
HANLEY2006	Rich	Not sure/not reported[1]	Convincing	Adequate	Yes	Clear
HERON2012	Rich	Reliable	Convincing	Adequate	Yes	Not sure/not reported[3]
HUNT2009	Rich	Reliable	Convincing	Adequate	Yes	Not sure/not reported[3]
MAPP2005A/2005B	Rich	Not sure/not reported[1]	Convincing	Adequate	Yes	Not sure/not reported[3]
MCCREIGHT2008	Rich	Not sure/not reported[1]	Convincing	Adequate	Yes	Not sure/not reported[3]
MCGRATH2013	Rich	Not sure/not reported[1]	Convincing	Adequate	Yes	Clear
NICHOLLS2007	Rich	Reliable	Convincing	Adequate	Yes	Not sure/not reported[3]
PARVIN2004	Rich	Not sure/not reported[1]	Convincing	Adequate	Not sure/not reported/not applicable[2]	Not sure/not reported[3]
PATEL2013	Rich	Not sure/not reported[1]	Convincing	Adequate	Yes	Clear
RAYMOND2009	Rich	Not sure/not reported[1]	Convincing	Adequate	Yes	Not sure/not reported[3]
ROBERTSON2003	Rich	Not sure/not reported[1]	Convincing	Adequate	Yes	Not sure/not reported[3]
RYNINKS2014	Rich	Reliable	Convincing	Adequate	Yes	Not sure/not reported[3]
SHAKESPEARE2003	Rich	Reliable	Convincing	Adequate	Yes	Not sure/not reported[3]
SHAKESPEARE2006	Rich	Reliable	Convincing	Adequate	Yes	Not sure/not reported[3]
SIMMONS2006	Rich	Reliable	Convincing	Adequate	Yes	Not sure/not reported[3]
SLADE2010	Rich	Not sure/not reported[1]	Convincing	Adequate	Yes	Not sure/not reported[3]

SMITH2007	Rich	Reliable	Convincing	Adequate	Yes	Not sure/not reported[3]
SNOWDON2012	Rich	Reliable	Convincing	Adequate	Yes	Clear
STANLEY2006	Rich	Not sure/not reported[1]	Convincing	Adequate	Yes	Not sure/not reported[3]
STAPLETON2008	Rich	Reliable	Convincing	Adequate	Yes	Not sure/not reported[3]
TEMPLETON2003	Rich	Not sure/not reported[1]	Convincing	Adequate	Yes	Not sure/not reported[3]
THOMSON2008	Rich	Reliable	Convincing	Adequate	Yes	Not sure/not reported[3]
THOMSON2013	Rich	Not sure/not reported[1]	Convincing	Adequate	Yes	Not sure/not reported[3]
THURTLE2003	Rich	Not sure/not reported[1]	Convincing	Adequate	Yes	Not sure/not reported[3]
TSARTSARA2002	Rich	Reliable	Convincing	Adequate	Yes	Clear
TURNER2008	Rich	Reliable	Convincing	Adequate	Yes	Not sure/not reported[3]
TURNER2010	Rich	Reliable	Convincing	Adequate	Yes	Not sure/not reported[3]
WITTKOWSKI2011	Rich	Reliable	Convincing	Adequate	Yes	Clear

Notes. [1] No double-coding is reported
[2] Ethical approval not reported
[3]The role of the researcher is not adequately described

Table 29: Matrix of qualitative evidence for service user experience

Dimensions of person-centred care	Key points on a pathway of care							
	Access	Information and support	Assessment and referral	Primary care	Therapeutic intervention	Assessment and referral to inpatient care	Hospital care	Discharge/ transfer of care
Involvement in decisions and respect for preferences	WITTKOWSKI 2011	ROBERTSON 2003 SHAKESPEARE 2006	COOKE2012 DEJONGE2001 EDGE2005/ 2007/2008 HALL2006 MCGRATH 2013	CHEWGRAHAM 2009 TURNER2008	BOATH2004 EDGE2011 HERON2012 MCGRATH 2013 SHAKESPEARE 2006 SLADE2010 TURNER2008 TURNER2010	-	ANTONYSAMY 2009 MAPP2005A/ 2005B NICHOLLS2007 SNOWDON2012 TEMPLETON 2003 THOMSON2008 THOMSON2013	HERON2012
Clear, comprehensible information and support for self-care	-	DEJONGE2001 HALL2006 HERON2012 MCGRATH 2013	-	-	-	-	NICHOLLS2007 SIMMONS2006 TSARTSARA2002	-
Emotional support, empathy and respect	CHEW-GRAHAM2009 EDGE2011	-	EDWARDS2005 HANLEY2006 MCGRATH 2013 PATEL2013	COOKE2012 SMITH2007 STANLEY2006 STAPLETON2008	BREUSTEDT 2013 SHAKESPEARE 2006 SMITH2007 TURNER2010	-	HUNT2009 MAPP2005A/ 2005B MCCREIGHT 2008 NICHOLLS2007	-

			SHAKESPEARE 2006				RYNINKS2014 SIMMONS2006 SNOWDON2012 THOMSON2008 THOMSON2013 TSARTSARA2002	
Fast access to reliable health advice	-	BOATH2004 HANLEY2006 SLADE2010	-	TEMPLETON 2003	-	-	ANTONYSAMY 2009 TSARTSARA2002	-
Effective treatment delivered by trusted professionals	AYERS2006 CHEW-GRAHAM2009 COOKE2012 DEJONGE2001 EDGE2005/ 2007/2008 EDGE2011 EDWARDS2005 HALL2006 HANLEY2006 MCGRATH2013 PARVIN2004 PATEL2013 RAYMOND2009 SHAKESPEARE 2006 SLADE2010 STANLEY2006 STAPLETON 2008 TEMPLETON 2003 THURTLE2003 TURNER2010 WITTKOWSKI	SMITH2007 TEMPLETON 2003 WITTKOWSKI 2011	EDGE2005/ 2007/2008 HALL2006 ROBERTSON 2003 SHAKESPEARE 2003 SHAKESPEARE 2006 SLADE2010 WITTKOWSKI 2011	CHEWGRAHAM 2009 HANLEY2006 SMITH2007 TEMPLETON 2003	AYERS2006 BOATH2004 EDGE2005/ 2007/2008 EDGE2011 HALL2006 HERON2012 MAPP2005A/ 2005B NICHOLLS2007 PATEL2013 RAYMOND 2009 ROBERTSON 2003 SHAKESPEARE 2006 SLADE2010 TEMPLETON 2003 THOMSON 2013 TURNER2008 WITTKOWSKI 2011	-	ROBERTSON 2003 SHAKESPEARE 2006	-

	2011							
Attention to physical and environmental needs	-	-	SHAKESPEARE 2003	-	COOKE2012 EDGE2011 RAYMOND 2009 SHAKESPEARE 2006 TURNER2010	-	ANTONYSAMY 2009 HERON2012 SIMMONS2006 TSARTSARA2002	-
Involvement of, and support for, family and carers	-	HERON2012	-	-	HERON2012 ROBERTSON 2003 THOMSON 2013	-	RYNINKS2014	-
Continuity of care and smooth transitions	HERON2012 SMITH2007	-	-	RAYMOND2009 STANLEY2006	BOATH2004 TURNER2008 TURNER2010	-	MAPP2005A/ 2005B NICHOLLS2007 RAYMOND2009	HERON2012

6.2.5 Summary of themes from the qualitative analysis of service user experience

Access

Key positive experiences

Continuity of care

Women highlighted the benefits of integrated identification and management for mental health problems, achieved through provision of care from a single known person or though collaboration between the professionals involved in their care.. Specifically, women who had experienced postpartum psychosis discussed how effective communication between healthcare professionals enabled them to focus on recovery and parenting (HERON2012):

> ... they had got a community nurse that would come out every week so she would assess how I was and I could talk to her about anything. And there were ups and downs, you know, there were times when I became really anxious and she got me in to see the psychiatrist earlier than my scheduled appointment on more than one occasion. (HERON2012, p. 160)

While, women who were being treated for substance misuse and had experienced a specialist home visiting service, were very positive about the provision of continual empathic support and access to specialist knowledge from a known person (SMITH2007):

> Just because you know that they're job ehm is working with that kind of thing so you know they accept like drug problems and its not really an issue I think. It's easier because you know its not an issue, it's easier to speak to people and get on with them and they're there to help you and that's why they're there. (SMITH2007, p. 26)

It is, however, important to note that although some women had positive experiences of integrated care, a recurring theme experienced across the care pathway was an unmet need for the sharing of information and treatment planning between professionals and a fragmented care plan.

Key negative experiences

Barriers to access

Women were frustrated that they could not access services unless they were in crisis (COOKE2012, EDWARDS2005, PATEL2013):

> I obviously needed some help... . I think there should be more awareness because if it took the doctor to come round twice, the midwife everyday and the paramedics to not even spot it, I just think its quite sad really that so

> many professionals couldn't spot it and I went to see an emergency doctor as well at NHS Direct ... so it was a bit of an ordeal to get me into hospital really, in the end it was my mum's doctor, the family doctor who came out after surgery to see me and he admitted me straight away because he knew I wasn't like that normally. (EDWARDS2005, p. 160)

> You shouldn't have to press that danger button of "I'm gonna self-harm" or "I'm gonna hurt my children" for someone to help you. (COOKE2012, p. 35)

Women experienced a number of barriers to accessing help from primary care, including system barriers such as difficulty in getting a GP appointment (CHEWGRAHAM2009) and experiences of GPs being unwilling to listen to, or dismissive of attempts to communicate, psychological distress (CHEWGRAHAM2009, RAYMOND2009, STANLEY2006):

> And I did actually mention something and my doctors were actually no use, they just turn around and said, 'oh well, it's the weather'. (STANLEY2006, p. 261)

> ... wouldn't go to the doctors because you can never get an appointment and it's crap. They always treat you like there's something else wrong and why are you wasting his time....I wouldn't have gone [to the doctors] even if I'd been dragged kicking and screaming (CHEWGRAHAM2009, p.5)

Women also felt that healthcare professionals were too busy to address psychological needs (EDGE2011, EDWARDS2005, STANLEY2006, TURNER2010, WITTKOWSKI2011):

> ... the health visitor said something like, 'you know in this community we have to look after a thousand and something babies' and that instilled in me the feeling like 'oh they are very busy these people and I don't have to be bothering them all the time'. So sometimes when you think of just calling them for something, you don't. (EDGE2011, p. 259)

Cultural differences were also perceived to create barriers to accessing help and support:

> In Pakistan we only saw lady professionals, but here you don't have a choice, you have to see the men as well otherwise you don't get to see a doctor. My husband is always at work so he can't come with me, I feel very uncomfortable. (WITTKOWSKI2011, p. 487)

> ... you need someone who's on the same wavelength as you, who shares the same cultural experiences as you, which sometimes isn't available... I wouldn't wanna particularly unburden myself to some White woman, if I'm honest about it. And that's the bottom line. It's about having someone who

> you can chat to who understands... where you're coming from... (EDGE2008, p. 385)

Moreover, the lack of information about services available could intensify feelings of isolation and desperation for an already vulnerable group of women (WITTKOWSKI2011):

> I need help and support zarroorat hey [desperately needed], my husband left me in pregnancy, and I have no-body, my family are in India. I can't speak English properly, and I can't read English to fill out forms. My GP says go the HV [health visitor] and HV says go to GP. I don't know what to do, I need help, don't know where to go, or who to turn to. (WITTKOWSKI2011, p. 486-487)

Barriers to disclosure

One of the most noticeable barriers to access experienced by women with mental health problems in pregnancy and the postnatal period, and a recurrent theme across the qualitative experience of care review, was that women felt reluctant to disclose difficulties to healthcare professionals for fear that their baby would be taken away from them (AYERS2006, COOKE2012, DEJONGE2001, EDGE2005/2007/2008, EDWARDS2005, HALL2006, HANLEY2006, MCGRATH2013):

> I spiralled into dark depression you know with all these horrible things that I was having to live with and too terrified to speak to anyone about for fear that they would take [the baby] away (AYERS2006, p. 393)

> So that's what really freaked me out about it, you know, like talking to the health visitor, because I don't want them to think that I'm not coping, and they might take my baby off me there. So I just tried to cope with it myself. (COOKE2012, p. 35)

> Concerns about stigma and fears of being perceived as a bad mother acted as barriers to self-referral (CHEWGRAHAM2009, RAYMOND2009, STANLEY2006, THURTLE2003, WITTKOWSKI2011):

> ...with my health visitor, I, I try not to, try not to let too much out because then she won't think I am a bad mum, if you see what I mean, so I tend not to let too much out with the health visitor. (CHEWGRAHAM2009, p. 5)

> I didn't want anyone to think I wasn't coping. (RAYMOND2009, p. 44)

> There is a huge stigma of being mentally ill in the public, but for us Asians there is a double disadvantage. I really fear that work will find out. (WITTKOWSKI2011, p. 487)

Women also described anxiety associated with their interactions with healthcare professionals where they felt that such interactions were dominated by risk assessment. Where women felt that risk assessments had been conducted covertly (for instance, professionals had not explained the reasons for taking detailed written notes), anxiety had been further increased (COOKE2012).

A lack of confidence in healthcare professionals was also described, with feelings that professional-service user interactions were formulaic and leaflet-driven (COOKE2012, EDGE2011, TEMPLETON2003):

> My experience has been: leaflet (baby massage); leaflet (postnatal depression); leaflet (baby immunisations). 'Any questions let us know. Any problems, [see your] GP'. It's leaflet, leaflet, leaflet; then 'see you later'. (EDGE2011, p. 259)

> Women were also not always sure about the role of the health visitor and the extent to which health visitors were responsible for their care or just for their babies (CHEWGRAHAM2009, SHAKESPEARE2006, SLADE2010), or just concerned with physical healthcare to the exclusion of the mental health problem (COOKE2012, PARVIN2004):

> It's not clear, you know [that she could help with postnatal depression]. I just look on her as the health visitor. If she'd said, you know, 'I'm trained and I can help you and I will sit and help you and I will listen to you and then I will tell the doctor what I think', then, yeah, I would have gone down to see her probably ... or asked her to come up here. (SHAKESPEARE2006, p. 159)

> I thought that the care would be more round care as opposed to just being about my baby's weight, which is basically all it's ever been about. (COOKE2012, p. 36)

> A related barrier to disclosure, and a recurrent theme, was the perception that healthcare professionals focused on the needs of the baby over the needs of the mother (EDGE2005/2007/2008, EDGE2011, RAYMOND2009, TURNER2010). For instance, women felt they had been treated like *a baby carrier* or *a walking womb* (RAYMOND2009, p. 45).

Women were also not hopeful that disclosure would lead to acceptable care and support (COOKE2012); for instance, they perceived antidepressants as the only treatment option available (EDGE2005/2007/2008, EDGE2011, TURNER2010):

> ... one of my friends got really depressed ... [her] GP offered her antidepressants and she refused ...all they are interested in is giving you drugs. They don't really give you social support. It's not about, 'what are

> your needs?' It's about 'how much can I drug you? Do you need sleeping tablets? Do you need antidepressants?' (EDGE2011, p. 260)

Information and support

Experience of information and support

Information and support provided through home visits

Women who were being treated for a substance misuse problem were very positive about the information provided to them by a specialist home visiting service, in particular, women described feeling supported and reassured by being informed about effects of drugs on the fetus and prepared for potential admission of their baby to the neonatal unit (SMITH2007):

> It was important that we have [specialist health visitor] as nobody else explained anything I needed to know about things, like if there were any side effects and [specialist health visitor] would tell you about different studies and just explained everything you needed explained both medical and everything else. (SMITH2007, p. 26)

However, it is important to note that the more representative experience of information and support for women with mental health problems during pregnancy or in the postnatal period was characterised by a number of unmet needs.

Unmet needs for general mental health information

Information to aid recognition

Women spoke about not knowing how to react when their symptoms (in this instance, of depression), did not disappear or increased in severity (HANLEY2006):

> I was frightened to tell anyone, but things had been getting on top of me. I thought it was just lack of sleep and this heavy cold. I thought that after a good night's sleep it would get better and I would be able to manage again. (HANLEY2006, p. 151)

Information about treatmentWomen also expressed a need for information tailored to their treatment or recovery stage and from other women. Women highlighted the importance of being spoken to directly and with respect for their agency even in circumstances where their capacity is impaired (HERON2012):

> I knew I was going to this Mother and Baby Unit whatever, it could have been mars for all I knew, but nobody was talking directly to me. As far as I understand it, I seemed able to understand everything going on around me, but my mind was in overdrive... Had somebody sat down and said: 'You've got this. You're going here. We're going to do this, that and the other. You'll be alright', maybe it wouldn't have been so bad. (HERON2012, p. 161)

> It's misleading information out there and I think we need to get proper advice out there to women to let them know you can get better... credible information, that was endorsed by, you know, the powers that be, to say that this is accurate and correct and it comes from those people who have looked into this illness the most, then you could trust that information. And go to that one place in the internet to find it all. (HERON2012. p. 161)

Age- and culturally-appropriate information and support

Teenage mothers spoke about their need for information about mental health and sources of support available, and also highlighted the importance of healthcare professionals being aware that teenage mothers might not be coping as well as they might pretend (DEJONGE2001).

Women from black and minority ethnic communities described information and support in the form of leaflets and insufficient face-to-face communication in pregnancy (TEMPLETON2003). South Asian women suggested a number of service improvements, including verbal and written information about depression in pregnancy, information about services available and culturally-specific support (WITTKOWSKI2011).

Unmet needs for post-diagnosis information and support

Post-diagnosis information about postpartum psychosis

Women described an unmet need for post-diagnosis information about postpartum psychosis. This was particularly important because they described needing to fill gaps in their memory with self-initiated information seeking (MCGRATH2013). Women with postpartum psychosis also highlighted a need for treatment information (ROBERTSON2003).

Post-diagnosis information about depression in the postnatal period

Women with symptoms of depression in the postnatal period described mixed experiences regarding post-diagnosis information about postnatal depression. Where information had been provided, women were positive (BOATH2004):

> They made me feel better about my postnatal depression because of them I fully understood what it was (BOATH2004, p. 228)

However, unmet needs for information and emotional support characterised the experiences of many women with depression in the postnatal period (HALL2006, SLADE2010):

> I didn't really know much about it to be honest ... nothing from a ... professional point of view. (SLADE2010, p. e444)

> It's really difficult to ask for help, whether it's the health visitor or the family. I didn't think there was any way they could understand. It is so hard to talk, to actually say the words. (HALL2006, p. 257)

Where post-diagnosis information about postnatal depression was lacking, women described the experience as confusing and wanted a discussion with their health visitor about the diagnosis and treatment options (SHAKESPEARE2006):

> No, no one tells you, no one tells you what they're thinking in their head, about, I wish people would do that, I mean, but she had some agenda in her head and she was going through it, she was thinking about it and she was poking, giving me questions but she didn't tell me what she was thinking about me and I want to know because I don't know what it is, you know, I don't, know what is it. (SHAKESPEARE2006, p. 158)

Unmet need for information and support for partner

Where information and support about postpartum psychosis was made available to their partner, women were very positive about the experience (HERON2012):

> I think it helped my husband first to be able to put a label on what was happening. Secondly, to realise that this is what happens in PP... I think it was reassuring for him to read about delusions and stuff, and to know that its quite common for women with PP to think they're the messiah or have special powers or you know. It was important to him in just seeing the process through ... to stick by me, to know that there was a treatment that could work... (HERON2012, p. 162)

However, in many cases women described an unmet need for information and support for their partners (HERON2012):

> My partner needed strategies to cope with the fear. Fear of relapse and fear of me not sleeping, or having another dip ... the ups and downs were just hideous for him... And also... because I did have two suicide attempts, and you know the fear for him of, 'what is she going to do next'. (HERON2012, p. 162)

> [Partners need] detailed but accessible information about what the condition is, that you're wife's going to recover, she's going to be 100% fine... She hasn't now turned into a basket case permanently, and she didn't mean what she said when she was horrible to you... (HERON2012, p. 162)

> It was hard for him. There wasn't much information out there... My husband I think was unsure whether he would ever get his wife back again. That's very distressing, when it doesn't need to be. (HERON2012, p. 162)

Assessment and referral

Barriers to disclosure in assessment

Stigma of diagnosis

Women talked about how the stigma of diagnosis could act as a barrier to disclosure in assessment because a 'label' was seen as a threat to their 'coping image', in terms of self-concept and in terms of the image women wanted to portray to healthcare professionals (COOKE2012, EDGE2005/2007/2008, SHAKESPEARE2006, SLADE2010):

> ...I don't want to be labelled...I don't want them to label me, they treat you differently and I think that makes you worse. I think you live to your label...if I think, 'I haven't got postnatal depression' and I don't want to do something, I can't blame it on my postnatal depression...if I start to label myself that I do [have postnatal depression], I can be very negative and I can't be bothered. Whereas once that option isn't there anymore [I say], 'come on, this isn't on', you know, I've got to find that piece of extra [strength] from somewhere and just get on and do it (EDGE2005, p. 21)

As a consequence of the perceived stigma attached to psychiatric diagnoses, women were reluctant to use the term 'depression' (EDGE2008, HALL2006):

> I was just embarrassed really. There's still a stigma to it, I thought postnatal depression, God they just kill their children, that's all you see in the media, y'know drama of they're going to kill all their children in a horrible nasty way and then be put away for the rest of their life. That's what postnatal depression was, and that's what I thought if I told people, they'd be like, better watch her. (HALL2006, p. 258)

Service user awareness

Another barrier to self-referral for assessment was women's lack of awareness about signs and symptoms of mental health problems (DEJONGE2001, EDGE2005/2007/2008), which rendered them reliant on healthcare professionals to translate their feelings into symptoms (EDGE2005/2007/2008):

> ...so I went to the GP and said, 'doctor, I just don't feel right'. 'I'm getting ill, I just don't feel right...what is it? (EDGE2008, p. 384)

Professional awareness

However, gaps in professional knowledge and awareness (EDGE2007, ROBERTSON2003), or unwillingness to recognise symptoms (EDGE2005/2007), could also compound women's feelings of fear and isolation:

> ...you have no idea what's going on, what's real and what's not, but when the doctors don't appear to know either that's really scary particularly when they're supposed to make you better (ROBERTSON2003, p. 419)

> He [GP] said, 'you're not depressed. Will you stop thinking you're depressed? I will send you for counselling if you want to go to counselling so you can talk, but you are not depressed'. (EDGE2007, p. 33)

Women suggested that early assessment and intervention would be a desired service improvement (WITTKOWSKI2011).

Fears about baby being taken away

In their interactions with primary care professionals, women said that they covered up feelings because they were afraid of losing their baby (HALL2006, SHAKESPEARE2003, SLADE2010):

> I didn't respond to the Edinburgh scale honestly... because I was scared what (the health visitor) would say. I was worried. I thought the baby would get taken off me. It wasn't until... I'd just had enough and I phoned up the health visitor. I said I need to see you, I think I need to be admitted into a psychiatric unit. (HALL2006, p. 257)

> I didn't trust them I suppose so I didn't tell the health visitors how I was feeling. (SHAKESPEARE2003, p. 618)

> I was so vulnerable, I believed what she [her mother] said, you know [about the baby being taken away]. (SHAKESPEARE2003, p. 618)

> I didn't want anyone's help to be honest after I had [my previous child]. I was so frightened that people would think I couldn't cope and take her off me. (SLADE2010, p. e443)

Professional-service user relationship

Some women found that their relationship with their health visitor hindered disclosure, either because they didn't emotionally engage with them or because they didn't know them well (SLADE2010):

> I did ask for support but I didn't really get any. And the health visitor's response ... 'Well you seem like you're doing alright', which kind of closes it off doesn't it then? (SLADE2010, p. e443)

> I didn't feel like talking to her. I didn't really know her that well so ... (SLADE2010, p. e443)

> .. So I think she wasn't as person-centred and she didn't really have the people skills to manage, you know, she could have, sort of offered advice

> and support in a much more supportive way instead of 'Well you haven't done this, you haven't done that', and her tone was all wrong as well. (SLADE2010, p. e443)

Experiences of diagnosis

Diagnosis reassuring

Women spoke about feelings of relief and reassurance upon being diagnosed (EDWARDS2005, HANLEY2006, MCGRATH2013, PATEL2013); for instance, one woman felt her condition had been *sanctioned* by her diagnostic label and other mothers spoke about the diagnosis giving them *permission to be ill* (HANLEY2006):

> Even though it was this thing you'd not heard of, it was a relief to know...it does exist, other people have had it before me and there are things that can be done. (MCGRATH2013, p. 6)

Stigma of diagnosis

However, a diagnosis was not reassuring to all women because a 'label' conferred stigma. Some women described how having a diagnosis meant that professionals tended to treat the label and not the person (MCGRATH2013). While for others being labelled with, for instance, postnatal depression was *scary* and something to be resisted (PATEL2013):

> ...but I was adamant that I was fine and that it was just a lack of sleep and this, that and the other and I would not let her refer me to anybody because I was fine, I was just blocking it out... (PATEL2013, p. 686)

Experiences of screening

In general, women described positive experiences of screening, as a shift of focus from baby to mother (SLADE2010).

Experiences of specific screening tools, of the EPDS in particular, were more mixed (SHAKESPEARE2003). Some women found that the closed question format made disclosure easier:

> I did think, gosh, this is good, because it's much easier to do this than to actually look somebody in the face and say, look, I am finding this really difficult to cope. Say look, discover me, please. (SHAKESPEARE2003, p. 616)

While for others closed questions were found to be restrictive:

> There's so much more that you want to say rather than just answering quite closed questions. (SHAKESPEARE2003, p. 616)

> If I was feeling bad, I'd rather have a coffee and a chat with someone, than put circles round numbers, while the baby's crying. (SHAKESPEARE2003, p. 616)

Some women found screening questions intrusive and frustrating in the absence of a solution.

The setting in which the EPDS was administered was also raised as an important factor contributing to women's experiences of screening, with some feeling that the baby clinic was an unsuitable environment for administration and stating a preference for screening at home:

> That first Edinburgh test, to have it filled in and then talked about in front of everybody else was just terrible. (SHAKESPEARE2003, p. 616)

Pre- and post-diagnosis information and support

Women highlighted that the lack of pre-diagnosis information about treatment options, or consequences of particular responses to questionnaires, resulted in a reluctance to complete the EPDS honestly (SHAKESPEARE2003):

> I was told this was a questionnaire to identify people having problems with postnatal depression and that was it, there was no treatment or no consequences discussed. It wasn't clear to me what would happen if I ticked the bad boxes. I should have been answering it for my own good, and people were trying to help me, but I wanted to get the answers right. (SHAKESPEARE2003, p. 616)

Women also expressed a need for post-diagnosis information and support; where feedback and information were provided after administration of the EPDS, the experience was valued. Women needed the health visitor to take time and be empathetic in talking about screening (SHAKESPEARE2003, SHAKESPEARE2006):

> And I was so grateful, and then I just talked to her, and it was so nice to be able to talk freely with her [about the EPDS] at the time. (SHAKESPEARE2003, p. 617)

> She [health visitor] said 'Oh dear, oh, that's not very good is it, oh, oh well, I, well we'd better, I'd better come and see you'. That's exactly what her sort of tone was, 'Naughty you' sort of thing. And I thought 'Oh, what have I done', you know, just the last person, you know, if I had, if I was feeling miserable or whatever, she's the last person in the whole wide world that would be of any help whatsoever, she's the most unsympathetic person and, you know, it has the opposite effect, makes you feel awful, you know. (SHAKESPEARE2006, p. 157-158)

Women emphasised the importance of follow-up after positive screening in particular (SHAKESPEARE2003):

> I purposely circled the things 'cos I'm struggling and it felt like the form was just left on the side and nobody picked it up and the health visitor didn't get back to me, which I'm really disappointed about, but I didn't have the courage to ring her up to ask her for help. (SHAKESPEARE2003, p. 617)

Primary care

Access to help and support

Information about available services

Women expressed a lack of awareness about the support available to them from primary care (TEMPLETON2003):

> I don't know what support is out there. (TEMPLETON2003, p. 214)

Continuity of care

Women spoke about the benefits of having support from a known professional in terms of facilitating access to services (RAYMOND2009, STANLEY2006):

> It was the not having to start explaining again to someone new which was so great. (RAYMOND2009, p. 45)

Women also expressed a need for a 'connection' with primary health care professionals in order to facilitate disclosure. Key components which women identified as being important to the development of professional-service user rapport were flexible boundaries, the perception of availability, respect, and empathy (COOKE2012):

> She goes if you need anything I'm always here, and she talked to me like a friend. (COOKE2012, p. 35)

Benefits of disclosure

Opportunities to raise distressing feelings were appreciated, and women felt that disclosure minimised feelings of isolation (STANLEY2006):

> They made me feel, they made me realise I wasn't on my own, that, all stuff that could be done ... (STANLEY2006, p. 261)

In addition to potential emotional support, women were also positive about the practical help and support offered by health visitors (HANLEY2006, SMITH2007, TEMPLETON2003).

Need for individualised help and support

A recurrent theme across women's experience of care was the need for individualised help and support, and the importance of avoiding a 'one size fits all' approach. This theme emerged as a general principle across the care pathway, but also in relation to specific information and support needs, which may vary across conditions and across service settings.

Treatment of the label not the person

Women who were receiving treatment for substance misuse problems described stigmatising interactions with their GP, where they felt that their individual needs were not listened to or addressed (SMITH2007):

> I just think that if I go and see him about a problem, even if it's just like [describing nature of problem] the first thing he'll ask me is about my drug problem and my methadone and that's not the issue and that's not why I'm going but everything is like linked to that and it's just I think that he looks down a little bit. (SMITH2007, p. 26)

Feeding support for women with an eating disorder

Another example of a specific need for individualised support was highlighted in the experiences of women with an eating disorder who required support for feeding their baby (STAPLETON2008). Women with an eating disorder described a lack of compassionate support for their feeding decision:

> I couldn't breastfeed. I just couldn't. I was desperate to get rid of the weight. I just wanted some reassurance from the midwives that bottle-feeding was all right but all they did was tell me off for not breastfeeding. (STAPLETON2008, p. 110)

> I know that yes, of course they've (midwives) got to encourage you to breastfeed, but they've also got to acknowledge that sometimes you just can't. I couldn't. I couldn't bear eating proper food anymore. (STAPLETON2008, p. 110)

Where personal support was received it was appreciated:

> One midwife was really nice. She said 'Don't be so stupid – my mother never (breast) fed me and I've got two degrees'. But the others tried to pressure. [...] All you want is that reassuring voice telling you it will be all right. (STAPLETON2008, p. 110)

The women's comments highlighted the potential for misinterpreting claims that breastfeeding helps weight loss. For instance, women expressed dissatisfaction if weight loss was not substantial or did not happen as fast as they had anticipated (STAPLETON2008).

Women reported problems with breastfeeding and/or with 'satisfying' the baby and expressed a need for information and support that was sensitive to their eating disorder:

> He'd just cry and cry but I couldn't satisfy him. He didn't seem to be getting enough from me. The health visitor told me to increase my fat intake to see if that would help. I felt really guilty but I couldn't do that. I'd put on so much weight in pregnancy already there was no way I could do that. (STAPLETON2008, p. 113)

> She (baby) started losing weight and I panicked. The health visitor came and said 'Get some Mars bars down you' – which of course I wasn't going to do. But it was just a glitch. It was just for a week where she didn't put weight on. I'm glad I didn't listen to the health visitor or I'd have been back into bingeing and vomiting. (STAPLETON2008, p. 113)

Treatment options

Women spoke about a reluctance to consult their GP because antidepressants were perceived as the only treatment option and regarded as unacceptable by some (CHEWGRAHAM2009, TURNER2008):

> That's all they have, GPs, and I just didn't want to go onto antidepressants, because obviously I've heard people get addicted to them and then you're stuck on them and you have a vicious circle (CHEWGRAHAM2009, p. 5)

However, other women were satisfied with antidepressants and GP care (HANLEY2006).

Therapeutic intervention

Unmet needs: specific intervention needs

Mother-baby relationship interventions

Mothers who had experienced a traumatic birth discussed problems with mother-baby attachment, including avoidant and over-protective feelings (AYERS2006, NICHOLLS2007):

> I could never just cuddle and hold her (AYERS2006, p. 395)

> I can remember thinking, you horrible thing, you've done this to me, and what you doing here, you evil child (AYERS2006, p. 395)

> I felt such a failure at actually giving birth that I was determined that I was going to do everything else (AYERS2006, p. 395)

> I was aware that I didn't have the feelings and I put on an act with [the baby]... I used to coo to her and all that sort of stuff but I didn't actually mean it... it was all fake, I honestly just did it because that's just what mothers are supposed to do ... (NICHOLLS2007, p. 502)

Mothers with symptoms of depression in the postnatal period expressed concerns around mother-baby attachment (HALL2006), including:

> I haven't bonded with my baby. (HALL2006, p. 257)

> I question if I really love my child. (HALL2006, p. 257)

Mothers who had experienced postpartum psychosis also expressed a need for help in learning how to interact with their babies (HERON2012):

> I wanted to learn stuff to do with my baby and for me that was massively missing. I invited over a health visitor and I asked 'please can you teach me how to interact with [my baby] 'cause I'm very depressed'. But I was terrified, absolutely terrified, that I wasn't doing the right things with her. I thought she wasn't gonna learn to talk or do anything because I wasn't interacting with her right. And the health visitor just didn't give me any practical tips at all... She was just saying 'you'll be fine', 'you'll get your confidence back' and dur-de-dur. I'm sure those all things were true, but tips, practical hands on tips. I really needed that. (HERON2012, p. 160)

Psychological treatment and support groups

There was a perceived need for psychological treatment (BOATH2004) and/or support groups (BOATH2004, EDGE2011, HERON2012, RAYMOND2009, ROBERTSON2003, WITTKOWSKI2011):

> Group therapy, yoga and individual counselling would have been nice to be offered and this could well of speeded a recovery being able to talk and be with others with similar problems (BOATH2004, p. 226)

> I think if I had had to get up to go to something it would have helped me to give the day a purpose, rather than sit around in my pyjamas. (RAYMOND2009, p. 45)

> If I'd have met people with similar experiences or could have had a conversation with somebody who'd been through the same thing...I didn't know of anyone at that time, so that would have been a big help. (HERON2012, p. 159)

> There should be someone there who could answer questions, maybe get the group going and then just the group could continue to meet..., so the women

> could get to talk freely amongst themselves about issues that are concerning them. (RAYMOND2009, p. 46)

In addition to peer support, women perceived the benefits of talking therapies and support groups to include the provision of: the security of regular support; structure to their day; an opportunity to escape their immediate surroundings (for instance, a small flat with no outside space); practical help and support; and the chance to educate and inform peers (RAYMOND2009).

Formal psychological support for partners

Women who had experienced postpartum psychosis spoke about the need for formal psychological support for their partner in order to address trauma and the breakdown of trust (HERON2012, ROBERTSON2003):

> ...trust is a big issue there, you know, a trust has been broken. They don't trust you because you have done all these strange things and you don't trust them because you think they will take you back to hospital. It's taken many, many, many months to solve. I feel if there was some system in place, where they could refer you to psychotherapy and the whole family would be involved so they can understand and you can understand them, it would definitely speed up recovery. (HERON2012, p. 162)

> ...the trauma of the memories cos I think for [my husband], he'd seen some of the pretty hideous stuff that I said and thought when I was so unwell, really quite dramatic things. He described it once to me as like a video playing over in his mind, and I think that's where you need someone who's a bit of a specialist to help, cos still if we talk or think about another baby, its that stuff that comes back. (HERON2012, p. 162)

Unmet needs: general principles of care

Interventions for the full spectrum of need

Women expressed an unmet need for care pathways that can provide support for the full spectrum of need from subthreshold symptoms to severe mental illness (EDGE2011).

Focused on needs of mother

Women also spoke about the need for a woman-centred approach (EDGE2011):

> ... somebody [is] not just checking on the baby but actually sitting down with you asking, 'how are you doing?' 'What can I do to help you?' (EDGE2011, p. 259)

Specialist treatment

Women with postpartum psychosis perceived themselves as different from people with other forms of mental health problems, because childbirth was the cause, and as such, they expressed a need for separate and specialist treatment (ROBERTSON2003):

> You're classed as a mental patient, rather than someone with an illness following childbirth, I think there's a difference you need specialist help (ROBERTSON2003, p. 419)

Professional-service user relationships

Women highlighted the need for trust, flexibility and responsiveness in the professional-service user relationship (MCGRATH2013):

> The very people you reach out to help you then become almost like your enemy, you're fighting against them and they're the people that were supposed to help us. (MCGRATH2013, p. 5)

Better follow-up

Women expressed a need for better follow-up care (BOATH2004):

> More care and better follow-up care from GP, midwife and health visitor. These people need to actually ask ''How are you'' rather than just assuming ... I would like better follow-up care (BOATH2004, p. 228)

Barriers to access: perception of interventions

Negative perception of antidepressants

Women expressed concern about taking antidepressants because they perceived these drugs to be addictive (CHEWGRAHAM2009, EDGE2007, TURNER2008) and sedative (EDGE2007, TURNER2008). Women were also concerned about the effects on their breastfed babies (EDGE2007, TURNER2008). Antidepressants were also regarded as stigmatising because there were implications that their problem was possibly severe (PATEL2013, SHAKESPEARE2006) or they were not coping (PATEL2013, TEMPLETON2003, TURNER2008):

> People will think she needs to be on meds to be a normal mother... (PATEL2013, p. 686)

> My concern is that I will just get addicted and it will change my personality (CHEWGRAHAM2009, p. 5)

> I approve of psychiatry, I approve of psychology, but I don't want to be a person who needs chemical adjustment. (SHAKESPEARE2006, p. 155)

> I didn't want it to become something really serious. You know, I didn't want the drugs, because I didn't want this to be serious depression, or ... you know, I wanted it to be something minor that would just, I wanted it to go. (SHAKESPEARE2006, p. 155)

The need for long-term monitoring, particularly in the context of the lack of continuity of care, also contributed to negative feelings about antidepressants (TURNER2008):

> I don't want to take tablets. I want to cope with it myself and then I don't have to go to the doctors every few minutes ... whenever I go, I don't ever see the same doctor, so every time I go I have to explain it all. (TURNER2008, p. 452)

Positive perception of antidepressants

Some women advocated the use of antidepressants but only if their mental health problem was severe or as a second-line treatment after non-response to psychosocial or psychological interventions (EDGE2011), or if they were in crisis or were waiting for psychosocial or psychological interventions (PATEL2013):

> I'd rather not, but it's the lesser of two evils I guess. (PATEL2013, p. 686)

Others felt that antidepressants were an acceptable first-line treatment, for instance, where social support was available (TURNER2008).

Perception of talking therapies

Women expressed mixed opinions regarding to the perceived efficacy of talking therapies (EDGE2007/2008):

> Counselling would make you a stronger person. You can't be strong on your own. (EDGE2008, p. 385)

> For some women it does work, like unburdening. For others, it doesn't. It's just reinforcing your life's crap (EDGE2007, p. 33)

Barriers to access: structural barriers

Waiting lists

Women talked about long waiting lists for counselling (EDGE2008).

Lack of childcare

Other structural barriers to visiting a counsellor included insufficient availability of childcare facilities (EDGE2008, TURNER2008):

> ...you have to have someone to look after your baby ... So who am I going to get to look after [baby]? You know, my family aren't here...she's being breastfed as well... (EDGE2008, p. 385)

> I did say was there any counselling that was available that I could access, and they said ''not really... (and) they don't come for you at home ...'' It was very difficult because I have two children to look after, in my present state of mind as well, like just driving a car and catching a bus is something that would be a nightmare for me. And they said the other option is antidepressants, and they started me on antidepressants. (TURNER2008, p. 453)

> Women also described feelings of being unable to leave the house and felt that, even if childcare was available, the social demands of attending clinical psychology clinics were too challenging given depleted self-confidence and lack of energy (COOKE2012). This led women to seek more accessible support through, for instance, internet chat rooms (COOKE2012):

> > Sometimes it kills me to just go school to drop [my son] off. (COOKE2012, p. 36)

Experiences of pharmacological intervention: antidepressants

Adherence

Women described how they self-regulated their antidepressant dosage, partly because of the stigma attached to its use (BOATH2004). Concerns about addiction also led women to wean themselves off medication (BOATH2004, TURNER2008):

> I take them only when I need them. (BOATH2004, p. 227)

> I do without when I can. (BOATH2004, p. 227)

Concerns about harms

Women were concerned about possible long-term effects of taking antidepressants (BOATH2004, PATEL2013, TURNER2008):

> I don't like taking tablets. They are bound to do you some harm in the long run. (BOATH2004, p. 227)

A good relationship with their GP was identified by women as an important factor in minimising concerns about antidepressants (TURNER2008).

Experiences of pharmacological intervention: antipsychotics

Involvement in treatment decisions

Women with postpartum psychosis discussed the need for greater consultation and negotiation in antipsychotic prescription, as they recognised the role of drugs in their recovery but felt that sedative effects interfered with their role as a mother (HERON2012):

> ... it would have been good I think to have been listened to about the side effects. I was on a very high dose of Olanzepine [sic] and it just knocks you out and makes you into a complete zombie... The psychiatrist was a young guy not understanding that we had needs as a family. My husband really needed me to be awake enough to get my baby dressed and you know, do that kind of stuff. It's just they're managing your risk of going high, maybe that's what they've got to do clinically, but I wanted a bit more of a human face of it really. (HERON2012, p. 159-160)

Women distinguished between clinical and social recovery and felt that while antipsychotics had addressed the former, they had negatively impacted upon the latter (HERON2012). Women also expressed a desire for follow-up counselling (HERON2012):

> When you're beginning to feel a bit better and you're not really seeing health professionals that much I think then, if you had—five or six sessions or something, with a counsellor and just went through how you felt about it. And you know, got a little bit of advice about how to cope with it. (HERON2012, p. 158-159)

Experiences of psychosocial interventions: listening visits and home visits

Professional-service user relationship

The experiences of listening visits or home visits appeared to be dependent on the quality of the relationship between the woman and the healthcare professional. Where women had a good rapport with their health visitor they were positive about listening or home visits. Components that contributed to positive professional-service user relationships included being knowledgeable about mental health issues, having time to listen and being empathetic and non-judgemental (SHAKESPEARE2006, SLADE2010, SMITH2007, TURNER2010):

> She [health visitor] was helpful ... to me in also being non-judgmental. I just find her ... I mean, there are just some people who you find are very comfortable to be with. (...). She's very good at seeing that you have time. I mean, she must be incredibly busy but she comes, she sits, she spreads, you know, you never feel like she's dying to go. (SHAKESPEARE2006, p. 156)

Conversely, a poor rapport was associated with negative experiences of listening visits, in particular, if the health visitor was perceived to be judgemental (SHAKESPEARE2006, SLADE2010):

> She [health visitor] came to see me and I felt like, I felt ... ten centimetres tall, all the time she was there. She, I don't know why, she didn't make me feel as though I was doing anything worthwhile at all. (SHAKESPEARE2006, p. 156)

Professional-service user relationship and settings for care

Inflexibility regarding settings for care could also compromise the relationship between the woman and health visitor (SHAKESPEARE2006):

> She wouldn't come here [to do the listening visits] cos she'd keep getting disturbed. My health centre's like a mile and a half down the road, and when you're not coping with a small baby and you've got to walk a mile and a half down the road, it's ridiculous. (SHAKESPEARE2006, p. 157)

Generally, home-based treatment was regarded positively because it provided privacy, comfort and the available facilities for entertaining and feeding their children, and alleviated the worry about going out and being late for an appointment (TURNER2010).

Need for individualised treatment

For some women the opportunity to talk to someone outside their family about how they were feeling was cathartic (SHAKESPEARE2006, SLADE2010, TURNER2010):

> I didn't have anyone to talk to and no one actually knew about me being diagnosed with postnatal depression, my mum or anyone, no one knew, not even my partner. So it was quite nice just to offload on someone. (TURNER2010, p. 236)

However, some viewed the non-directive approach as too narrow a model for a long-term approach (SHAKESPEARE2006, SLADE2010):

> Yeah, I think it was a catharsis type of thing, I mean the first time, I felt better after the first talk, and then the next one I felt was a bit annoying and then the next one I got a bit more annoyed with it, I just didn't know what the point was. I didn't see a purpose and she didn't explain it clearly. In the end she, I think she felt the same way, she wanted to be done with it, so, so it was sort of mutual. (SHAKESPEARE2006, p. 160)

Length of intervention

Some women considered eight visits insufficient to address their postnatal depression. As a consequence, women described feeling *left hanging* and *completely exposed* at the end of treatment (TURNER2010):

> Just me thinking about it [the idea of no treatment after the visits] now makes me feel quite panicky... what would have been the point of ripping off the plaster and starting to abrade the wound, only to then just say, oh well. (TURNER2010, p. 237)

Experiences of psychosocial interventions: support groups

Benefits of peer support

Women were positive about the opportunities to meet other women and discuss shared experiences, which support groups offered (HANLEY2006, TEMPLETON2003):

> Each week I look forward to going. It sounds crazy really but it is the only time I get to meet adults of like mind! (HANLEY2006, p. 151)

Women also viewed support groups as an opportunity to educate and inform peers (HERON2012):

> I joined a postnatal depression and illness support forum, and told my whole story on there, actually its funny 'cause I'm reflecting on it now, three years down the line and I think it was helpful at the time because I, just had this really strong need to educate and inform other people about it, you know?... I felt that I was almost making sense of the experience that had happened to me by educating others. (HERON2012, p. 158)

However, an unmet need for multicultural group support was highlighted (EDGE2011, TEMPLETON2003).

Social vulnerability

Conversely negative feelings towards support groups were expressed by some women who felt that group situations were not useful during early recovery (HERON2012):

> ...with support groups, if you're still feeling vulnerable you don't really want to go and expose yourself with other people, so it's much better to have something where you can get information and get support, without having to feel vulnerable like that. (HERON2012, p. 159)

Experiences of psychosocial interventions: interventions for traumatic birth

Benefits of post-birth discussion

Women were positive about the opportunities for discussion and debriefing following a traumatic birth (MAPP2005A/2005B, THOMSON2013):

> He took us all the way through it and we were able to ask questions. He answered our questions fully and honestly, which we were very grateful for. We found that crucial in our understanding with fitting things together and in accepting it. (MAPP2005B, p. 37)

> ...she put me in touch with X [Consultant Midwife] which is just the best thing that could ever have happened. Going through it (traumatic birth) really put my mind straight about a lot of things... (THOMSON2013, p. 768)
> ... we came out of that meeting [after birth services] and we felt we were on the road to recovery (THOMSON2013, p. 769)

Benefits of partner involvement

Women were also positive about the involvement of their birth partner in post-traumatic birth discussions, as an opportunity for women and their partners to share each other's version of events (THOMSON2013).

Hospital care

General experiences of hospital care

Lack of continuity of care

Women spoke about how fragmented healthcare made it more difficult for them to discuss their feelings with healthcare professionals (RAYMOND2009):

> Every time I went to see the midwife, or..., I always had somebody different, and I don't want to tell 10 people my story. (RAYMOND2009, p. 45)

Language barriers and lack of communication

Women from black and minority ethnic groups talked about negative experiences of hospital care, specifically language barriers and not being told what was happening to them (TEMPLETON2003).

Experiences of mother and baby units

Security and being with their baby

Women preferred being admitted to the MBU, rather than a general psychiatric ward, because they felt safer and believed that having their baby with them aided recovery (ANTONYSAMY2009).

Professional-service user relationship

Women were positive about their communication with healthcare professionals in the MBU (ANTONYSAMY2009):

> Sometimes people think you haven't got a brain and there's no point explaining to you. But the doctor here explained to me everything and I appreciate that (ANTONYSAMY2009, p. 360)

> The nurses are good. I can't think of anything else (ANTONYSAMY2009, p. 360)

Unmet needs

Access was raised as an issue in relation to a lack of local provision of MBUs (SHAKESPEARE2006). Where they were available, women discussed a need for improved access to doctors and nurses within the unit (ANTONYSAMY2009), and they also spoke negatively about the lack of organised ward activities (ANTONYSAMY2009).

Experiences of general psychiatric units

Being with the baby

Women experienced distress and anger at being separated from their baby on admission to a general psychiatric ward and talked about how this negatively impacted upon their confidence in resuming the mothering role after discharge (HERON2012).

Unmet need for specialist treatment

Women who had experienced postpartum psychosis expressed frustration and anger over the lack of specialist treatment available to them in a general psychiatric unit (HERON2012, ROBERTSON2003):

> I think being sent to what I feel was the wrong environment really made things worse, because there was no, sort of, specialist help or treatment in the psychiatric hospital. My partner wasn't able to stay with me, and I wasn't able to have my baby with me either. I was there for about 3 weeks. Eventually they let my baby stay with me once I'd got a bit better, but again, being in that environment wasn't good for either of us. There was somebody doing cartwheels and there was somebody throwing themselves on the floor... (HERON2012, p. 159)

> I was given treatment that everybody else on the ward had, nobody I saw had specialist knowledge of puerperal psychosis (ROBERTSON2003, p. 419)

Experiences of post-miscarriage or post-stillbirth information and support

Emotional support, empathy and respect

Women highlighted the need for professionals to recognise that miscarriage or stillbirth is traumatic and not routine (MCCREIGHT2008, SIMMONS2006):

> Most people treat miscarriage as not very important ''everybody has them'' etc. but it was very traumatic for me. (SIMMONS2006, p. 1,942)

Women also found the medicalising language used by healthcare professionals in relation to miscarriage distressing (MCCREIGHT2008, SIMMONS2006):

> My miscarriage was a 'missed abortion' type—(I hate this term for a wanted baby) (SIMMONS2006, p. 1,942)

> [one woman described her response to the term 'spontaneous abortion'] *I felt the doctor was implying that I had had an abortion and that I was to blame.* (MCCREIGHT2008, p. 9)

> Women who had experienced a stillbirth or miscarriage described a notable lack of empathy demonstrated by healthcare professionals during their interactions and treatment (MCCREIGHT2008):

> Before I had the anaesthetic I couldn't stop crying and the anaesthetist said 'could you stop crying, you're not the first, you won't be the last, my wife's had four of these.' And I asked him if they could take my baby out in one piece and he said 'if it comes out in one piece, it comes out in one piece'. (MCCREIGHT2008, p. 10)

> I was pregnant again when I went to see him (psychiatrist) and having concerns that this baby might also die. He told me that his wife had just had a baby and they were being kept awake all night, and I would soon know all about once this baby was born. (MCCREIGHT2008, p. 10)

Settings for care

Women who had just experienced, or were in the process of experiencing, a miscarriage described the negative impact of being cared for in an inappropriate setting (SIMMONS2006, TSARTSARA2002):

> I was admitted to a mixed ward with women who were still pregnant, women who were having voluntary terminations. I was admitted at 10 a.m., operated on at 7 p.m. I found the whole experience appalling. The concern seemed only to be for my physical well being, emotionally this was completely the wrong environment. In the morning I discharged myself and

> walked home a matter of a few hundred yards. I was offered no formal support. (SIMMONS2006, p. 1,942)

> I was very, very tearful and I think it's because you go down [to the antenatal clinic] and you go through all these seats of women who are about 8 months pregnant, 5 months pregnant. And you know that you've lost the baby, and you have to wait there, I think I waited about an hour to get my scan done. And it seemed, it seemed very very upsetting, a very poor system to me ... And I don t like jumping queues, but I think that is a very good cause to go straight to the front of the queue. (TSARTSARA2002, p. 59)

Unmet need for post-miscarriage information and follow-up support

Women expressed a need for clear and comprehensible information about the processes of miscarriage so as to alleviate distress (SIMMONS2006, TSARTSARA2002):

> It would have been valuable to have received information about what could happen and what to do, as I was at home when I lost the baby and it was an extremely distressing experience. (SIMMONS2006, p. 1,942)

Women described the follow-up support available as 'patchy' and suggested improvements included a simple follow-up check-up, bereavement counselling or a miscarriage group (SIMMONS2006) or a home visit from a midwife (TSARTSARA2002).

Positive experiences of specialised miscarriage units

Women spoke positively about the provision of individualised treatment and the perception of continuous accessibility and availability offered by a specialised miscarriage unit (TSARTSARA2002):

> There were loads offered to me. I mean they asked me if I wanted a counsellor... they were just really kind. And she said to me 'look, I know it's an early pregnancy, but even that, at the end of the day I could tell you wanted the baby'. They were really nice. And she said, 'even if after, perhaps sort of 6 months, you still find that you would like to talk to somebody, get in touch with us and we'll arrange something'. (TSARTSARA2002, p. 59)

Experiences of traumatic birth

Lack of control

In describing their experiences of a traumatic birth, women discussed distress associated with a lack of control over events (MAPP2005A/2005B, NICHOLLS2007, SNOWDON2012, THOMSON2008, THOMSON2013):

> Being awake in theatre doesn't help because you are in their domain and it is definitely their domain and they do what is easiest to save your life but the care of the mind is not looked at, at all. (MAPP2005A, p. 33)

> Nobody said to me, 'Is this alright? do you mind five or six complete strangers having a look at the most intimate parts of your body, sitting there with your legs in the air and the whole thing on display?' (NICHOLLS2007, p. 496)

> I wasn't involved with it (childbirth) because all my requests were met with a no (THOMSON2008, p. 271)

> ...even though they're around you, it's like you're just an object (THOMSON2013, p. 767)

Related to this lack of control, women discussed negative experiences of physical restraint during labour (NICHOLLS2007):

> They told [my husband] to come in and then got [my husband] to pull me upright, [midwife] on one arm and [my husband] on the other ...which I think was actually a terrible thing to do because it sort of brought an element of violence and restraint into our relationship which had not obviously been there before. And I was just fighting to get down. (NICHOLLS2007, p. 496-497)

It is, however, important to note that some women were satisfied with clinical decisions being made on their behalf during a crisis (MAPP2005A/2005B, SNOWDON2012):

> I was in their hands and let them carry on with it. I knew they had to do what was best. (MAPP2005A, p. 33)

Inadequate and/or inaccurate information

Where information was given during (MAPP2005A/2005B) or after (SNOWDON2012) a traumatic birth it was valued:

> The midwife was talking to me which did help, I felt as if there was a safety net there. (MAPP2005A, p. 32)

> [A]s I came round they must've been telling me over and over the same thing all the time...[I]t must've been going in because when they were talking to me when I was kind of, you know, conscious, I felt like I already knew most of it....Obviously they were being very brief, that I'd gone back to theatre again and I'm in intensive care, I'd had lost a lot of blood and I'd still

> got my uterus and the baby's fine. And they [put] a photograph of the baby...in my hand. (SNOWDON2012, p. 795)

Women discussed the need to be given information about what was happening during birth (NICHOLLS2007) and described a lack of communication during crises and after childbirth (MAPP2005A/2005B, SNOWDON2012):

> Being informed of what was happening in layman terms would have actually taken a lot of the stress and worry away and the panic, definitely the panic. (MAPP2005B, p. 37)

> I can't talk now but I'll talk to you later, can be helpful, because at least you'll get that sense of feeling that somebody wants to talk, but they are very busy at the moment. (MAPP2005B, p. 37)

> ...nobody said anything – at all. I think the consultant said, good morning, and that was it. The rest of the time he talked to the other doctors, no one talked to me. I wasn't there. (NICHOLLS2007, p. 498)

> [N]urses were just coming in, rushing in from God knows where, I mean I don't know how many there was and it felt like no one was telling me what was going on. I mean I was just lying there thinking 'Oh God, oh God, what's happening?' I suppose 'cos they were so concerned that I was bleeding so much... [T]hey were putting like stuff in me hands and...because they wasn't talking to me, I was worried, I was panicking. (SNOWDON2012, p. 793)

Longer term effects of lack of post-traumatic birth discussions

Women talked about how a continued lack of understanding about the traumatic birth could be 'a big problem' (MAPP2005A/2005B, SNOWDON2012):

> I was never debriefed properly. I don't know what happened during them days... It was all coping with the trauma and coping with the new baby...it probably took me till about six to eight months to actually come up with some of these questions that I wanted answers to, that Jerry couldn't answer 'cos obviously he didn't know the technicalities of it. So I feel like I've been left quite ignorant ... To this day I don't know what's happened. (SNOWDON2012, p. 796)

Focus on babies over mothers

Women described how they felt excluded from decisions during a traumatic birth because the focus was on the baby rather than them (THOMSON2013):

> ...she [midwife] said something along the lines of 'I'm not thinking about you now I'm thinking about this baby, that baby's my patient' as if saying you're going to have to let me do this'. And I couldn't argue with that. Alright I'd read a few books, but I'd never seen a labour or had experience of labour and I could not stand my ground in the face of somebody saying well I've got to think about this baby (THOMSON2013, p. 767)

Professional-service user relationship

Women talked about the need for compassionate care and to have their preferences taken into account (NICHOLLS2007, THOMSON2008):

> The people who are there to help you should be making it better not worse...the attitude of the people, the way they treat you, and pain relief. I think, you know, if those two things had been handled differently I would have had a totally different experience ... if they'd been handled differently...I don't think I would have ended up with PTSD. (NICHOLLS2007, p. 498)

> It was a male doctor, um, I have a history of depression and anxiety and I don't like being touched. I have very clear personal boundaries, and a male doctor came in, and I was like 'I can cope, It's only a doctor, It's only an examination, I can cope', and I just lay down on the bed, I just, melt down, started to cry, couldn't cope. [My husband] said to the guy 'stop' and he was like, 'well I've started it now' ... then it continued. (NICHOLLS2007, p. 498)

Continuity of care

Continuity of care and seeing familiar faces was viewed positively (MAPP2005A/2005B). However, more commonly, women emphasised a lack of communication between professionals during a traumatic birth (NICHOLLS2007, THOMSON2008):

Every person that came in, I had to give them my medical history because they didn't know, there didn't seem to be any hand over happening (NICHOLLS2007, p. 498)

Experience of stillbirth or termination of pregnancy following diagnosis of fetal abnormalities

Seeing and/or holding the dead baby

Women described how they were encouraged by midwives to see their dead baby following termination of a pregnancy (because of fetal abnormalities) and that they were motivated to make this decision because they wanted visual reassurance that something was wrong (HUNT2009):

I wanted to see the lesion on his spine because I wanted to be absolutely sure that there had been no mistake (HUNT2009, p. 1,114)

Women who had experienced a stillbirth described mixed feelings upon seeing their baby. For some women, the opportunity to see their baby, and to compare the baby's appearance to family members engendered feelings of relief (RYNINKS2014):

Her feet, they were like her dad's, she had big toes (laughs) it was just the fact she was so perfectly formed, all the creases on her hands and feet, and the nails and the hair starting to come through and stuff like that (RYNINKS2014, p. 6)

Holding her, seeing what she looked like, knowing whether she looked like me or like (partner). This might sound strange but I wondered if she'd have a crossover toe like me but she didn't. Her hair was like her dad's, dark and curly. You pin all your hopes on what they'll be like and I feel robbed of it. If I hadn't seen her it'd be 10 times worse as I'd never have known her. I can be at peace knowing that I'd held her. I needed that. (RYNINKS2014, p. 5)

Women also spoke positively about the experience of seeing and/or holding their stillborn baby in the context of the opportunity to form memories of the baby (RYNINKS2014):

It was (reassuring), and it wasn't what I expected at all and it was fine...nice in a way because we've got no other memories apart from me being pregnant and feeling her move inside me, we've got nothing else at all because she didn't breathe, she didn't have a life, so to have those memories is quite nice really. (RYNINKS2014, p. 6)

It was just being able to say goodbye to her properly, getting memories and things to remember her by, and just having cuddles and things. It was a special time. (RYNINKS2014, p. 5)

Conversely, some women (whose baby's body had been damaged or deteriorated) found the physical appearance of their baby disturbing and struggled with seeing or holding their baby (RYNINKS2014):

Unfortunately because she'd been inside me for some time and it was a pretty horrible forceps delivery in the end, had a bit of a problem in getting her out, a lot of the skin had come off so all down her side there was no skin and some of her arms and her face um and (partner) found that quite difficult. So when I was bathing her it was like 'I don't know how you can do that, I don't know how you can do it'. (RYNINKS2014, p. 6)

Women perceived the seeing and holding of their stillborn baby as initiating a process of acceptance of their loss. As such, this was either resisted because the women were still in a state of disbelief and were not ready to deal with their feelings, or was appreciated as a way of coping with the loss and accepting that their baby had died (RYNINKS2014):

I didn't want to hold him, and I think that was almost upholding the illusion that he was alive in this basket, and if I held him it would be obvious that he wasn't alive, and looking at him in the basket it was like he was asleep. (RYNINKS2014, p. 7)

I got to say goodbye to him, that he was my baby, whether he was alive or dead. That everyone got to see him. Got to touch him. (RYNINKS2014, p. 7)

It helped me to realise that she was dead. I think had we not seen her, err, it was a very, very real thing to have a dead body with you and yeah she's dead, you know what else could she be, here she is, and if I hadn't had seen her I'd be thinking 'well is the doctor telling me the truth, is she dead, is somebody kidnapped her and bringing her up somewhere else' you know that was all it as well. Umm, yeah I had forgotten that actually, I did think that at the time that it was quite important to see her. (RYNINKS2014, p. 7)

Women described a varying sense of satisfaction or regret with their decisions regarding seeing or holding their baby (RYNINKS2014):

I wouldn't have done anything differently um I definitely would have seen her. And I guess I almost can't believe I didn't want to, it would have been quite hard not to have seen her. It definitely helped... I think I would have felt worse now if I hadn't, you can't take that back, you can't go backwards and change it, so I definitely think it was the right thing to do and I guess I'm quite grateful for, I mean it wasn't, it wasn't pushy, but it was recommended (RYNINKS2014, p. 7)

I do I regret not holding him, and I think I regret not holding him purely because I never held him. Now, you know, I do regret not holding him. I think I should have been braver, but it's very easy to say that in hindsight. Cause at the time couldn't so. And maybe I was right at that time, cause if I had of held him I would have actually felt that physical sense of not having my baby in my arms. So perhaps it was a sort of self-preservation defence mechanism kicking in. (RYNINKS2014, p. 7)

Spending time with the dead baby

Women who had experienced a stillbirth described the opportunity to spend time with their baby as a cathartic experience (RYNINKS2014):

> It was quite nice to have that time with her, looking back on it now. Even thinking about it at the time... Yes, it was so horrendous and so heart breaking, I'm glad we did it and spent time with her. (RYNINKS2014, p. 4)

Involvement of partners and family

For women who had experienced a stillbirth, opportunities for their partners and family to be involved in the protocols following stillbirth (for instance, to also be given the opportunity to see and hold the stillborn baby) were appreciated (RYNINKS2014):

> ...Important everyone else got to see him because they are so close to me, and they were so close to me throughout the pregnancy as well. And they are excited about it. Yeah. Yeah I just wanted them to see how real he was. I wanted to make sure that anyone who wanted to hold him had held him. (RYNINKS2014, p. 4)

> They dressed him. (Partners) parents came over to be with us. When (partner) and I were together we really dwelled. When other people were there we chatted about other stuff. My mum and dad were in the delivery suite waiting. (Partners) mum wanted to see him, dad wasn't sure. We didn't want to put pressure on them, they had to do it for themselves, then it was all of us together. It was nice that all of them came and they shared that with us. It's a shared experience. (RYNINKS2014, p. 5)

Mementoes

Mixed opinions and experiences of mementoes following termination of a pregnancy because of fetal abnormality were described. Some women described how photographs or mementoes were taken of the baby by hospital staff as a matter of course and how they appreciated the time this allowed them to make the decision about whether or not to see and keep these photographs or mementoes (HUNT2009):

> They said to us, 'We've taken a footprint and a handprint' ... I thought it was really nice that they did actually do these things, because I've subsequently read in people's, other people's experiences, and they say they wish they had seen the baby, they wish they had asked for footprints and things. And it's quite nice to know that they're there and if, if, you know you don't want them at first, maybe after a period of reflection you would want that. (HUNT2009, p. 1,117)

> ...we had read, and we're really glad we did, the SATFA booklet at the time, and that says, you know, it said, ''You may want to see the baby, hold the baby, have photographs''. And we didn't take a camera with us. We felt that, it seemed morbid. So we actually asked, and they were of course incredibly busy and we had to keep asking for the photograph. They offered us, I think it was probably hospital policy to offer handprints and footprints because obviously they'd be used to dealing with stillbirth... . I remember at the time

> we had to be quite persistent to get our photograph, which isn't very nice, but I'm glad we have it. And certainly the handprints and footprints, I'm very glad we have those... for years at a time we haven't looked at them, but we know they're there ...it is a comfort to know they're (HUNT2009, p. 1,117)

While others found questions about commemorating the baby and the experience of photographs being taken of their baby upsetting (HUNT2009):

> When I went to the postnatal check they gave me all the photographs that had been taken in the hospital. I had the polaroids, but I was given a film of photos of my baby. And I actually really wished they hadn't, they hadn't done that ... I wasn't really expecting it. The doctor that I saw spoke to me in a very hushed voice like somebody was dead in the next room which made me feel quite uncomfortable. And then all these photographs arrived and I remember sitting there in the consulting room by myself looking at all these photographs of this baby and it just triggered something in my head. (HUNT2009, p. 1,118)

> I was very definite that I didn't want photographs, because to me that's just it's, it's the moment of death, I don't want to see him dead baby, I just don't. (HUNT2009, p. 1,118)

> 'Would you like a little Moses basket with sort of white covers on?' And 'Would you like us to take hand and footprints?', and all this sort of thing. And that really upset me quite a bit, because I didn't want to think of it as a baby. I, it was just a dreadful mistake, something gone horribly wrong, and I wanted to get out of there really. And all this talk about hand and footprints was really quite upsetting (HUNT2009, p. 1,118)

Preparation and the importance of individualised treatment

The mixed experiences of seeing and holding the baby and of keeping mementoes following a termination of a pregnancy because of fetal abnormalities or a stillbirth highlights the importance of individualised treatment. Women expressed a desire to be provided with information and support to prepare them for making a decision about whether to see and/or hold the dead baby (HUNT2009, RYNINKS2014) and for decisions about a funeral (HUNT2009):

> I guess having some time and then seeing her was quite good. You feel like you're, you're coming to a bit more. I think if we'd have seen her too soon after I wouldn't have been really quite with it enough. (RYNINKS2014, p. 5)

> It was preparing for what was he going to look like, were we going to feel a bond with him, or were we going to feel disgust, we were worried and concerned about that. (RYNINKS2014, p. 5)

Discharge/transfer of care

Unmet needs

Support for hospital-home transition

Women who were being transferred from psychiatric inpatient care to care in the community described the hospital-to-home transition as challenging because of low self-esteem and lack of confidence in their mothering skills. This unmet need left women feeling isolated and unsupported (HERON2012):

> ... because of the anxiety I was suffering after it, that, like I say, wasn't me at all, I didn't want to be left on my own. And the transition from 24 hour care for eight weeks to suddenly having nothing really, other than my husband's bit of time off work, but being self dependent again was for me, the hardest part of those six months after coming out... (HERON2012, p. 160)

> ...eventually I begged them to let me go home, and I wasn't really well enough when I was at home and there wasn't really an awful lot of support after I went home. (HERON2012, p. 160)

Suggested improvements

Home-based post-discharge support

Women with postpartum psychosis suggested that home-based one-to-one support from a healthcare professional with expert knowledge of postpartum psychosis who could give practical advice on caring for the baby, would be beneficial in order to support the hospital-to-home transition (HERON2012):

> I saw my psychiatrist once every two weeks to check on my medication. It would have been good to have somebody who knew something about it, like a sort of social worker or community mental health worker or something, to visit and just ... give you some help and encouragement. I mean that's why it's great if they can come to your home because, as somebody who has been to visit psychiatrists quite a lot in their offices, it's quite daunting and you tend to, especially as a female, you're always eager to please and 'oh I'm doing fine' and put your best face on it. (HERON2012, p. 160)

6.2.6 Summary of evidence from the primary qualitative review

Based on the review of the qualitative evidence for the experience of care for women with a mental health problem in pregnancy or the postnatal period, the following common themes were found to resonate across the care pathway:

- unmet need for collaboration between professionals and continuity of care
- stigma and fears about losing their baby acting as a barrier to disclosure
- healthcare professionals perceived as too busy or unwilling to address psychological needs
- focus on babies over mothers

- importance of non-judgemental and compassionate support from healthcare professionals
- importance of service user involvement in treatment decisions and individualised treatment
- need for longer-term follow-up and support.

6.3 LINKING EVIDENCE TO RECOMMENDATIONS

Taking into account the recommendations in *Service User Experience in Adult Mental Health* (NICE, 2011d; NCCMH, 2012) and *Patient Experience in Adult NHS Services* (NICE, 2012b; National Clinical Guideline Centre, 2012), the GDG determined that recommendations for this guideline should be specific to women with a mental health problem in pregnancy and the postnatal period, and should not replicate recommendations already covered in other NICE guidance. The GDG also agreed that some of the themes that emerged from the review of the experience of care (see Section 6.2.5) would be more appropriately addressed in other chapters of the guideline. Therefore the evidence from this review supports the development of recommendations in three separate areas of the guideline: (1) recommendations that are concerned with improving the experience and effectiveness of recognition and assessment (see Chapter 5); (2) recommendations for treatment (see Chapter 7 and 8); (3) and recommendations relating to all other aspects of care for a mental health problem in pregnancy and the postnatal period, including discussion and decision-making about treatment options, communication and information giving, and coordination of care.

The GDG was of the view that the review of a range of well-conducted primary studies was both comprehensive and of high quality. In addition the themes that emerged were in line with the experience reported by service user members of the guideline and also the concerns about women's experience of care expressed by clinical and academic members of the GDG.

In reviewing women's experience for this guideline, the GDG was concerned about both the lack of information given to women and the point in their care at which the information was provided. The consequences of this are various and include the decision by 90% of pregnant women to stop psychotropic medication when they discover they are going to have a baby. The GDG therefore saw the importance of developing a recommendation on providing information about mental health problems to all women of childbearing potential, which covers use of contraception, ascertaining whether the woman plans to become pregnant, the ways in which pregnancy and childbirth might affect a mental health problem, and the ways in which a mental health problem and its treatment might affect the woman and her fetus or baby and have an impact on parenting. For women who are already pregnant or in the first postnatal year, the GDG wished to ensure that culturally relevant information is given to all women (and if she agrees, her partner, family or carer) about mental health problems in pregnancy and the postnatal period. Furthermore, in order to address some of the barriers to accessing care that can be attributed to stigma, the GDG was keen to ensure that women understand that

mental health problems are not uncommon at these times and that healthcare professionals should foster hope and optimism about treatment.

A key problem identified in *Service User Experience in Adult Mental Health* was the lack of engagement of service users in decisions about their care. The review undertaken in this chapter confirmed that this was also the experience of women with a mental health problem in pregnancy and the postnatal period. In addition the review highlighted that women may also feel reluctant to talk about their problems out of a fear and a perception that healthcare professionals will form a negative impression of them as competent mothers. The GDG was conscious of the sensitivities that arise from this and also the impact on other family members, and was keen to ensure that the woman's role in caring for her baby was acknowledged and reinforced in a non-judgemental and compassionate manner.

The GDG was also concerned about problems with inter-professional communication and organisation, especially between professionals working in different agencies (for example mental health and maternity services), which emerged from the review of the experience of care. The GDG therefore advocated an integrated care plan for all women with a mental health problem in pregnancy and the postnatal period that outlines the care and treatment for the mental health problem and the roles of all healthcare professionals involved including who is responsible for coordinating the integrated care plan, the schedule of monitoring, and providing the interventions and agreeing the outcomes. The GDG also considered it important that the healthcare professional coordinating the integrated care plan takes responsibility for ensuring that everyone involved in a woman's care is aware of their responsibilities, there is effective sharing of information among services and with the woman herself, that mental health and wellbeing should be taken into account as part of all care plans, and that all interventions for mental health problems are delivered in a timely manner taking into account the stage of pregnancy or age of the baby.

The evidence relating to young women (teenagers) came from one study, and echoed the need for information about mental health problems in pregnancy and the postnatal period expressed by adult women in other studies. The GDG was keen, however, to make a recommendation for this age group, given the particular challenges relating to issues of consent and confidentiality, and therefore saw no reason to remove the recommendation from the previous guideline. In addition, the GDG wished to highlight the recommendations for young pregnant women made in the NICE guidance on *Pregnancy and Complex Social Factors* (NICE, 2010) that seek to address barriers to accessing and engaging with antenatal care for this vulnerable group. The GDG were also particularly concerned that young women may fall through gaps between services during the transition from adolescent to adult services for the assessment, treatment and management of their mental health problem, and as such a recommendation was made to ensure continuity of care during this transfer.

Finally, while the GDG was concerned not to replicate the recommendations from *Service User Experience in Adult Mental Health* (NICE, 2011d; NCCMH, 2012) and *Patient Experience in Adult NHS Services* (NICE, 2012a; National Clinical Guideline Centre, 2012), they thought it important to draw attention to the recommendations in both those guidelines. This was, in part, to emphasise that much of the experience of a mental health problem is common to all people with a mental health problem irrespective of whether or not they are pregnant or have given birth.

6.4 RECOMMENDATIONS

Using this guideline in conjunction with other NICE guidelines

Improving the experience of care

6.4.1.1 Use this guideline in conjunction with the guidance on service user experience in adult mental health (NICE clinical guidance 136) and patient experience in adult NHS services (NICE clinical guidance 138) to improve the experience of care for women with a mental health problem in pregnancy or the postnatal period. **[new 2014]**

Considerations for women of childbearing potential

6.4.1.2 Discuss with all women of childbearing potential who have a new, existing or past mental health problem:

- the use of contraception and any plans for a pregnancy
- how pregnancy and childbirth might affect a mental health problem, including the risk of relapse
- how a mental health problem and its treatment might affect the woman, the fetus and baby
- how a mental health problem and its treatment might affect parenting. **[new 2014]**

Principles of care in pregnancy and the postnatal period

Supporting women and their partners, families and carers

6.4.1.3 Acknowledge the woman's role in caring for her baby and support her to do this in a non-judgmental and compassionate way. **[new 2014]**

6.4.1.4 Involve the woman and, if she agrees, her partner, family or carer, in all decisions about her care and the care of her baby. **[new 2014]**

6.4.1.5 When working with girls and young women with a mental health problem in pregnancy or the postnatal period:

- be familiar with local and national guidelines on confidentiality and the rights of the child
- be aware of the recommendations in section 1.4 of the guideline on pregnancy and complex social factors (NICE clinical guideline 110)
- ensure continuity of care for the mental health problem if care is transferred from adolescent to adult services. **[new 2014]**

Coordinated care

6.4.1.6 Develop an integrated care plan for a woman with a mental health problem in pregnancy and the postnatal period that sets out:

- the care and treatment for the mental health problem
- the roles of all healthcare professionals, including who is responsible for:
    - coordinating the integrated care plan

- the schedule of monitoring
- providing the interventions and agreeing the outcomes with the woman. **[new 2014]**

6.4.1.7 The healthcare professional responsible for coordinating the integrated care plan should ensure that:

- everyone involved in a woman's care is aware of their responsibilities
- there is effective sharing of information with all services involved and with the woman herself
- mental health (including mental wellbeing) is taken into account as part of all care plans
- all interventions for mental health problems are delivered in a timely manner, taking into account the stage of the pregnancy or age of the baby. **[new 2014]**

Treatment decisions, advice and monitoring for women who are planning a pregnancy, pregnant or in the postnatal period

Information and advice

6.4.1.8 Provide culturally relevant information on mental health problems in pregnancy and the postnatal period to the woman and, if she agrees, her partner, family or carer. Ensure that the woman understands that mental health problems are not uncommon during these periods and instil hope about treatment. **[new 2014]**

6.4.1.9 Discuss treatment and prevention options and any particular concerns the woman has about the pregnancy or the fetus or baby. Provide information to the woman and, if she agrees, her partner, family or carer, about:

- the potential benefits of psychological interventions and psychotropic medication
- the possible consequences of no treatment
- the possible harms associated with treatment
- what might happen if treatment is changed or stopped, particularly if psychotropic medication is stopped abruptly. **[new 2014]**

6.4.1.10 If needed, seek more detailed advice about the possible risks of mental health problems or the benefits and harms of treatment in pregnancy and the postnatal period from a secondary mental health service (preferably a specialist perinatal mental health service). This might include advice on the risks and possible harms of taking psychotropic medication while breastfeeding and how medication might affect a woman's ability to care for her baby (for example, sedation). **[new 2014]**

6.4.1.11 Mental health professionals providing detailed advice about the possible risks of mental health problems or the benefits and harms of treatment in pregnancy and the postnatal period should include discussion of the following, depending on individual circumstances:

- the uncertainty about the benefits, risks and harms of treatments for mental health problems in pregnancy and the postnatal period
- the likely benefits of each treatment, taking into account the severity of the mental health problem
- the woman's response to any previous treatment
- the background risk of harm to the woman and the fetus or baby associated with the mental health problem and the risk to mental health and parenting associated with no treatment
- the possibility of the sudden onset of symptoms of mental health problems in pregnancy and the postnatal period, particularly in the first few weeks after childbirth (for example, in bipolar disorder)
- the risks or harms to the woman and the fetus or baby associated with each treatment option
- the need for prompt treatment because of the potential effect of an untreated mental health problem on the fetus or baby
- the risk or harms to the woman and the fetus or baby associated with stopping or changing a treatment. **[new 2014]**

6.4.1.12 When discussing likely benefits and risks of treatment with the woman and, if she agrees, her partner, family or carer:

- acknowledge the woman's central role in reaching a decision about her treatment and that the role of the professional is to inform that decision with balanced and up-to-date information and advice
- use absolute values based on a common denominator (that is, numbers out of 100 or 1000)
- acknowledge and describe, if possible, the uncertainty around any estimate of risk, harm or benefit
- use high-quality decision aids in a variety of numerical and pictorial formats that focus on a personalised view of the risks and benefits, in line with the guidance on patient experience in adult NHS services (NICE clinical guideline 138)
- consider providing records of the consultation, in a variety of visual, verbal or audio formats.**[new 2014]**

7 PSYCHOLOGICAL AND PSYCHOSOCIAL INTERVENTIONS FOR THE PREVENTION OR TREATMENT OF MENTAL HEALTH PROBLEMS

7.1 INTRODUCTION

Pregnancy, childbirth and the first postnatal year is a unique period of change for women. This period of transition may interact with women's psychological, social and biological vulnerabilities, culminating in psychological distress and mental ill health. The effects of poor mental health during the perinatal period can be especially difficult for women during a time when they face additional expectations and infant care demands. Further, emotional distress and problems during pregnancy, childbirth and the postnatal period warrant particular attention because of the longitudinal impact these difficulties have on the developing fetus and newborn baby, effects which are often mediated through the woman's disrupted relationship with her infant.

Psychological difficulties in pregnancy and the postnatal period range from minor transient disturbance with rapid unaided adjustment through common mental health problems to severe psychiatric disturbance. Pregnancy, childbirth and the demands and transitions associated with having a new child may precipitate or worsen psychological problems or lead a woman to seek help for previous and/or long-standing difficulties at this time.

Given that the nature of most mental health problems in pregnancy is little different from that of mental health problems of non-pregnant women in both their presentation and course, it is reasonable to assume, in the absence of evidence to the contrary, that treatment developed for non-pregnant women is likely to be effective. However, a number of factors specific to pregnancy and the postnatal period may alter the efficacy of psychological treatments in pregnancy and the following postnatal year. These include access, both in terms of the availability of the treatments and the women's capacity (relative to increased physical demands and childcare demands), the relative cost effectiveness of the treatments and, in particular, the need to consider the relative benefits of drug and psychological treatments in light of the increased risk of harm to the fetus associated with pharmacological treatment in pregnancy or during breastfeeding.

This chapter is concerned with reviewing psychological and psychosocial interventions for the prevention or treatment of mental health problems in the

pregnancy and the postnatal period, together with health economics evidence where appropriate. It also considers broader psychosocial interventions, such as protocols for mothers whose babies are stillborn.

7.2 FACTORS TO CONSIDER IN THE EVALUATION OF PSYCHOLOGICAL AND PSYCHOSOCIAL TREATMENT

7.2.1 Prevention versus treatment distinction

There is a great deal of inconsistency across studies in how disorders in pregnancy or the postnatal period are characterized, for instance, psychiatric diagnosis compared with scoring above a threshold on a scale (clinician-rated or self-report). This variability is also reflected in how researchers define their trials as preventative or as treatment. This lack of consistency makes it difficult to assess like for like within meta-analyses. Therefore, for the purposes of clarity and transparency it was decided that this review would use inclusion criteria and/or baseline mean symptom scores to make the distinction between prevention and treatment studies. Where participants in a trial had a psychiatric diagnosis the study was included in the treatment review. However, where the disordered group were defined based on symptomatology, consistent criteria (Table 30) were used to categorise subthreshold symptoms and symptoms of the disorder into the treatment review and below threshold symptoms into the prevention review. It is important to note that these cut-offs are distinct from symptomatology as an outcome, in which case we are limited by the thresholds selected by the trials and these are frequently higher (with moderate rather than mild cut-offs).

Table 30: Criteria for categorising prevention and treatment studies

Scale	Prevention	Treatment: Subthreshold	Treatment: Symptoms
BDI	<9	9-10	>10
BDI-II	<13	13-14	>14
Center for Epidemiologic Studies Depression Scale (CES-D)	<15	15-16	>16
EPDS	<8	8-9	>9
Hamilton Rating Scale for Depression (HRSD)	<7	7-8	>8
Hospital Anxiety and Depression Scale (HADS)	<7	7-8	>8
Impact of Events Scale (IES)	<34	34-35	>35
Quick Inventory of Depressive Symptoms (QIDS)	<5	5-6	>6

State-Trait Anxiety Inventory-State (STAI-S)	<39	39-40	>40
Wijma Delivery Expectancy Questionnaire (W-DEQ-A)	NA	NA	≥100

7.2.2 Review strategy and sub-analyses

The review strategy was to evaluate the clinical effectiveness of the interventions using meta-analysis by intervention. Following this, sub-analysis was conducted (dependent on available data), based on: risk factor for prevention studies (risk factors identified) or baseline diagnostic status for treatment studies (clinical diagnosis [usually assessed using structured psychiatric interview]; symptoms [above a pre-specified threshold on a rating scale]; subthreshold symptoms [just below a pre-specified threshold on a rating scale]); treatment timing (antenatal and/or postnatal); mode of delivery (for instance, face-to-face, internet, telephone and so on), format (individual and/or group), and intensity (low [<8 sessions contact with a healthcare professional]; moderate [8-15 sessions of contact]; high [≥16 sessions of contact]).

7.3 DEFINITIONS OF PSYCHOLOGICAL AND PSYCHOSOCIAL INTERVENTIONS

This chapter considers non-pharmacological treatments, including psychological therapies such as CBT and IPT and psychosocial interventions such as social support. The definitions of the main psychological and psychosocial treatments covered in this guideline are listed below.

7.3.1 Cognitive behavioural therapy

CBT for depression was developed by Aaron Beck during the 1960s. One of the assumptions underlying this form of therapy is that psychological distress is strongly influenced by patterns of thinking, beliefs and behaviour. Depressed patients have patterns of thinking and reasoning that focus on a negative view of the world (including themselves and other people) and what they can expect from it. Psychological distress may be alleviated by altering these thought patterns and behaviours without the need to understand how earlier life events or circumstances may have contributed to how those patterns arose. A key aspect of the therapy is an educative approach, where the patient learns to recognise their negative thinking patterns and how to re-evaluate them. The new approach needs to be practised outside of the sessions in the form of homework.

CBT is a discrete, time-limited, structured psychological treatment. The patient and therapist work collaboratively to identify the types of thoughts, beliefs and interpretations and their effects on current symptoms, feeling states and problem areas. The patient then develops the skills to identify, monitor and counteract problematic thoughts, beliefs and interpretations related to the target symptoms. The

patient also learns a repertoire of coping skills appropriate to targeting thoughts, beliefs or problem areas. CBT is usually delivered as an individually focused therapy but has also been developed as a group treatment. Common antenatal and postnatal modifications include delivery in the home of the mother or mother-to-be.

7.3.2 Co-parenting intervention

This intervention is based on the assumption that the postnatal period may be a time of increased stress not just in terms of the transition to motherhood but also in terms of marital adjustment as women attempt to handle both maternal and marital roles. The intervention involves partners in therapy sessions, and positive interaction and communication between the couple is encouraged by discussing strategies for child care and housework.

7.3.3 Directive counselling

This intervention incorporated elements of supportive listening and history taking in common with listening visits (non-directive counselling) but also included more directive techniques of problem clarification, goal formation, problem solving and partner sessions. This intervention can be delivered individually or in a group format.

7.3.4 Home visits

A structured series of prenatal and infancy visits by either lay home visitors or health professionals to provide emotional and practical support (such as how to care for the infant and/or how to access appropriate health and social services).

Home visitors can assist parents to improve: the outcomes of pregnancy, by helping women improve their prenatal health; children's subsequent health and development by helping parents provide competent infant and toddler care; maternal physical and mental health by facilitating access to appropriate community services; mother–infant interactions by helping mothers to be sensitive and respond to their child's behavioural cues; parents' economic self-sufficiency by helping them complete their education, find work, and plan future pregnancies.

7.3.5 Infant sleep interventions

Infant sleep interventions such as controlled crying and camping out, are based on behavioural principles. Controlled crying describes the process of sleep training whereby parents respond to their infant's cry at increasing time intervals, and is based on the principle that infants need to be taught to fall asleep independently in order to self-settle after night waking. Camping out is based on the same underlying principles as controlled crying but involves a parent sitting with their infant until they fall asleep and gradually removing their presence over a few weeks. These interventions involve the provision of information about normal sleep cycles and the development and management of sleep problems, and discussion and development of individually tailored sleep-management plans.

7.3.6 Interpersonal psychotherapy

IPT was developed by Klerman and colleagues (1984) initially for depression, although its use has been extended to other areas (Weissman et al., 2000). It may be defined as a discrete, time-limited, structured psychological treatment derived from an interpersonal model of affective disorders that focuses on interpersonal issues. The patient and therapist work collaboratively to identify effects of key problem areas related to interpersonal conflicts, role transitions, grief and loss, and social skills, and their effect on current symptoms, feeling states and/or problems. The treatment seeks to reduce symptoms by learning to cope with or resolve these interpersonal issues.

IPT focuses on current relationships and interpersonal processes and on the difficulties that arise in the daily experience of maintaining relationships and resolving difficulties. The main tasks are to help patients to link their mood with their interpersonal contacts, recognising that, by appropriately addressing interpersonal problems, they may improve both relationship and mood. There is usually an agreed focus for treatment, such as interpersonal role transitions. Therapy sessions concentrate on facilitating understanding of recent events in interpersonal terms and exploring alternative ways of handling interpersonal situations. IPT is usually delivered as an individually focused therapy but has also been developed as a group treatment. Common antenatal and postnatal modifications include delivery in the home of the mother or mother-to-be.

7.3.7 Listening visits (non-directive counselling)

Counselling was developed by Rogers (1957) who believed that people had the means for self-healing, problem resolution and growth if the right conditions could be created. These include the provision of positive regard, genuineness and empathy. Rogers' original model was developed into structured counselling approaches by both Truax and Carkhuff (1967) and Egan (1990). Voluntary sector counselling training tends to draw on these models. Counsellors are trained to listen and reflect patient feelings and meaning (Rogers, 1957). Many other therapies use these basic ingredients of client-centred counselling, but there are differences in how they are used. Holden and colleagues (1989) developed the concept of 'listening visits' based on these Rogerian, non-directive counselling skills and this has been taken up by a number of healthcare professionals working in the postnatal area, in particular health visitors. The healthcare professional is trained to help clients to gain better understanding of their circumstances and themselves. The therapist adopts an empathic and non-judgemental approach, listening rather than directing but offering non-verbal encouragement, reflecting back to assist the person in making decisions. This approach is usually offered by briefly trained healthcare professionals rather than mental health professionals and often takes place in the client's home.

7.3.8 Mindfulness training

Mindfulness-based cognitive therapy was developed with a specific focus on preventing relapse/recurrence of depression (Segal et al., 2002). It is derived from mindfulness-based stress reduction and CBT for acute depression. Mindfulness-based cognitive therapy is intended to enable people to learn to become more aware of the bodily sensations, thoughts and feelings associated with depressive relapse, and to relate constructively to these experiences. It is based on theoretical and empirical work demonstrating that depressive relapse is associated with the reinstatement of automatic modes of thinking, feeling and behaving that are counter-productive in contributing to and maintaining depressive relapse and recurrence (for example, self-critical thinking and avoidance) (Lau et al., 2004). Participants learn to recognise these 'automatic pilot' modes, step out of them and respond in healthier ways by intentionally moving into a mode in which they 'de-centre' from negative thoughts and feelings (for example, by learning that 'thoughts are not facts'), accept difficulties using a stance of self-compassion and use bodily awareness to ground and transform experience. Common postnatal-specific modifications include the presence of babies in the room during sessions and replacing a longer single meditation per session with a few shorter meditations.

7.3.9 Mother-infant relationship interventions

Mother-infant relationship interventions are psychological interventions where the goal is to improve the relationship between the mother and infant. These interventions are based on a psychological theory about the nature of attachment between the mother and infant. These interventions typically involve observations of mother–infant interactions, feedback (often video-based), modelling and cognitive restructuring. The primary aim is to enhance maternal sensitivity to child behavioural cues and awareness of the child's developing skills and needs.

7.3.10 Music therapy during delivery

This intervention involves listening to self-selected music during spontaneous vaginal delivery. The intervention is based on the principle that music may have anxiolytic and analgesic properties and improved satisfaction with the childbirth experience is also hypothesized to impact upon depression in the postnatal period.

7.3.11 Non-mental health-focused education and support

A structured educational treatment (often offered in groups) which may focus on preparation for childbirth (antenatal/in pregnancy) or practical aspects of childcare (postnatal). Such interventions offer an integrated approach to pregnancy, delivery and the mental and physical health and well-being of the woman and the infant and may include a focus on the social and personal adjustment to the role of a parent following the birth of a child (Gagnon, 2000).

7.3.12 Peer-mediated support and support groups

Peer-mediated support is a system of giving and receiving help founded on key principles of respect, shared responsibility, and mutual agreement of what is helpful and is primarily in one direction with a clearly defined peer supporter and recipient of support. Peer volunteers who are mothers themselves and also have a history of antenatal or postnatal mental health problems are recruited and trained to deliver interventions. These interventions can include befriending and mentoring.

Support groups also provide an opportunity for peer support but are usually facilitated by a healthcare professional and discussions are usually structured around a series of pre-defined topic areas (for instance, transition to motherhood, postnatal stress management, co-parenting challenges). However, the primary goal of these interventions is to enable mutual support by bringing women into contact with other women who are having similar experiences and providing opportunities for sharing problems and solutions.

7.3.13 Post-miscarriage interventions

Post-miscarriage interventions may take the form of self-help, facilitated self-help or counselling, all with the common aim of providing meaning to the miscarriage experience. Intervention content typically includes discussion of: coming to terms with the loss; sharing the loss; resuming life as a non-pregnant woman; trying again.

7.3.14 Post-traumatic birth discussion and/or counselling

The purpose of the intervention is to: explain to women what happened in delivery; give the woman an option to discuss labour, birth, and post-delivery experiences; and to answer any questions she has. The content of the discussion is determined by each woman's experiences and concerns and the intervention is delivered by midwives and obstetricians who are experienced in talking with women about birth, able to listen with empathy to women's accounts, and aware of the common concerns and issues arising. It is important to note that this intervention does not include post-trauma debriefing (based on adapted Critical Incident Stress Debriefing [Mitchell, 1983]).

7.3.15 Pre-delivery discussion and psychoeducation

This intervention is aimed at addressing tokophobia (fear of childbirth) and typically involves the provision of information about childbirth and an opportunity to discuss previous obstetric experiences, feelings and misconceptions. This psychoeducative discussion can be delivered individually or in a group format. Such discussions may be psychologically-informed, for instance, incorporating CBT principles of focusing on the target problem and reformulation of this problem through self-reflection and cognitive restructuring, and may also include guided relaxation exercises.

7.3.16 Protocols for women following stillbirth

Protocols for women following stillbirth may include seeing and/or holding the stillborn infant, keeping photographs or mementoes and having a funeral.

7.3.17 Psychologically (CBT or IPT)-informed psychoeducation

Psychoeducation is a structured educational treatment (often offered in groups), which may focus on preparation for childbirth (antenatal) or practical aspects of childcare (postnatal) but also includes a specific mental health component with information about common mental health disorders in the antenatal and/or postnatal period. These interventions are often informed by psychological principles and as such techniques from CBT and/or IPT are used such as cognitive restructuring, pleasant event scheduling, role play, guided relaxation, and homework exercises. The research on psychologically-informed psychoeducation interventions has most commonly involved women with subthreshold symptoms of depression, but has also been used for women with subthreshold symptoms of OCD.

7.3.18 Psychosomatic interventions

These interventions involve a comprehensive psychosomatic assessment, supportive therapy, psychoeducation and relaxation techniques and are guided by the principle that stress associated with pregnancy may be linked to the long-term course of anxiety, depression and physical complaints.

7.3.19 Self-help and facilitated self-help

Self-help interventions are psychological interventions typically based on cognitive behavioural principles that seek to equip people with strategies and techniques to begin to overcome and manage their psychological difficulties. Self-help usually provides information in the form of books or other written materials that include psychoeducation about the problem and describe techniques to overcome it. Although computerised interventions have the potential to be interactive and individualised, those that have been tested in clinical trials are, for the most part, relatively fixed programmes. In 'pure' self-help, only the written materials are used, in facilitated self-help, a therapist or alternatively a computer-based system (stand alone or web based) assists the service user in using the materials.

7.4 PSYCHOLOGICAL AND PSYCHOSOCIAL INTERVENTIONS FOR THE PREVENTION OF MENTAL HEALTH PROBLEMS

7.4.1 Introduction (prevention)

Prevention of disease is the ultimate quest for all working in healthcare but is rarely achievable, particularly in complex human conditions such as mental health problems. Antenatal and postnatal mental health care offers tantalizing theoretical opportunities for prevention, not just in this generation but the next and beyond. In common with most preventative health care, primary prevention in the field of

antenatal and postnatal mental health presents the greatest challenge and is likely to rely on interventions outside the traditional remit of health services. For example, a recent study found that the strongest predictor of antenatal depression was the woman's own history of childhood maltreatment (Plant et al., 2013).

It is in secondary prevention (limiting the development or recurrence of mental health problems) and tertiary prevention (reducing the effects of mental health problems on mother and child) that antenatal and postnatal mental health care offers unique and realistic opportunities as we have advanced notice of periods of known high risk, in identifiable high risk groups, amongst a population that has universal contact with health professionals. Furthermore, current evidence suggests that the potential target outcomes are not restricted to mental disorders in the mother, but could extend to physical health, exposure to maltreatment and intellectual and social functioning in the child. However, evidence on the effectiveness of preventative interventions is only just beginning to emerge and is at present meagre, although some important conclusions are possible. These have led to both positive and negative recommendations of relevance to service planners, clinicians and women themselves. Nevertheless, it is striking that important clinical dilemmas remain uninformed by robust trial evidence.

7.4.2 Clinical review protocol (prevention)

The review protocol summary, including the review question(s) and the eligibility criteria used for this section of the guideline, can be found in Table 31. A complete list of review questions can be found in Appendix 8; further information about the search strategy can be found in Appendix 10; the full review protocols can be found in Appendix 9.

The review strategy was to evaluate the clinical effectiveness of the interventions using meta-analysis. However, in the absence of adequate data, the available evidence was synthesised using narrative methods. An analysis of all interventions was conducted and graded. Following this sub-analysis was conducted (dependent on available data), based on risk factor, treatment timing, format (individual and/or group), and intensity. Where possible both an available case analysis and an intention-to-treat (ITT) analysis (worst case scenario [WCS]) were used.

Table 31: Clinical review protocol summary for the review of psychological and psychosocial interventions for the prevention of mental health problems

Component	Description
Review question(s)	RQ 2.1 What is the effectiveness of selective preventative interventions (for women with no risk factors) in reducing the likelihood of developing mental health problems in pregnancy or the postnatal period? RQ 2.2 What is the effectiveness of indicated preventative interventions (for women with identified risk factors present) in reducing the likelihood of developing mental health problems in pregnancy or the postnatal period?

	RQ 2.3 What strategies should be adopted to minimise potential harm to the women or the fetus/infant of these interventions?
Population	**Included** **Review question 2.1** Women who are pregnant or in the postnatal period (from delivery to the end of the first year). Inclusion is not based on any other baseline risk factors. **Review question 2.2** Women who are pregnant or in the postnatal period (from delivery to the end of the first year) who are considered to be 'at risk' of developing mental health problems. Include women: • with a history of a mental health problem but who do not meet diagnostic criteria for mental health problems at the current time • experiencing major life events • with a family history of mental health problems • with psychosocial risk factors (for example, SES) • who have infants with regulatory problems • who experienced an operative delivery or traumatic birth • who experienced a pre-term delivery (<37 weeks gestation) and/or whose infant had a low birth weight • who experienced a miscarriage • who are adolescents • experiencing intimate partner violence Exclude women: • who are currently receiving treatment (psychosocial or pharmacological) for an existing mental health problem (see review of interventions for the treatment of a mental health problem) • who are not pregnant or in the postnatal period (up to 1 year postnatal)

Intervention(s)	**Included interventions**
	• Psychosocial or psychological interventions for women with no pre-specified baseline risk factors (other than being pregnant or in the postnatal period) (RQ 2.1) or for women with at least one identified baseline risk factor (RQ 2.2), including:
	  • Home visits
	  • Peer-mediated support and support groups
	  • Post-traumatic birth counselling
	  • Psychologically (CBT or IPT)-informed psychoeducation (booklet or group)
	  • Mother-infant relationship interventions
	  • Non-mental health-focused education and support
	
	Excluded Interventions
	• Universal prevention programmes (that is, targeted to the general public or to a whole population group that has not been identified on the basis of increased risk [here increased

	risk was defined as being pregnant or in the postnatal period])
Comparison	**Review question 2.1 and 2.2** • Treatment as usual, enhanced treatment as usual, no treatment, waitlist control • Another active prevention intervention
Critical outcomes	**Maternal Outcomes** • Symptom-based • Diagnosis of mental disorder • Symptomatology (clinician- and self-report) • Relapse • Service utilisation o Hospitalisation for mental health problems o Retention in services (assessed through drop-out rates as a proxy measure) • Experience of care o Satisfaction o Acceptability of treatment (including drop-out as a proxy measure) • Quality of life o Quality of life measures o Functional disability o Social functioning o Social support o Perceived parenting stress • Harm o Side effects (including drop-out because of side effects) • Quality of mother–infant interaction and infant care o Quality of mother–infant interaction measures (including maternal sensitivity and child responsivity) o Establishing or continuing breastfeeding **Fetal/Infant outcomes** • Fetal and infant physical development (including congenital malformations) • Side effects • Cognitive development of the infant • Physical development of the infant • Emotional development of the infant • Optimal care of infant (for example, vaccinations, well-baby check-ups) • Prevention of neglect or abuse of the infant • Service use o Planned (health visitor, vaccinations, well-baby check-ups) o Unplanned (emergency department visits, inpatient, urgent or acute care) o Social service involvement
Study design	**Review question 2.1 and 2.2** Systematic reviews of RCTs Primary RCTs For the review of protocols following stillbirth cohort and case-control study designs were included **Review question 2.3**

	N/A; GDG consensus-based

7.4.3 Studies considered[11] (prevention: identified risk factors)

Twenty-two RCTs reported across 25 papers met the eligibility criteria for this review: ARACENA2009 (Aracena et al., 2009), BARLOW2007 (Barlow et al., 2007), BARNET2007 (Barnet et al., 2007), BRUGHA2000 (Brugha et al., 2000), COOPER2009 (Cooper et al., 2009), EASTERBROOKS2013 (Easterbrooks et al., 2013), GORMAN1997/DENNIS2013 (Gorman, 1997; paper unavailable, so data extracted from Dennis & Dowswell, 2013), HARRIS2006/DENNIS2013 (Harris et al., 2006; paper unavailable, so data extracted from Dennis & Dowswell, 2013), HOWELL2012 (Howell et al., 2012), KERSTING2013 (Kersting et al., 2013), KIEFFER2013 (Kieffer et al., 2013), MEIJSSEN2010A/2010B/2011 (one study reported across three papers: Meijssen et al., 2010a; Meijssen et al., 2010b; Meijssen et al., 2011), MELNYK2006 (Melnyk et al., 2006), MEYER1994 (Meyer et al., 1994), NEWNHAM2009 (Newnham et al., 2009), PHIPPS2013 (Phipps et al., 2013), RAVN2012 (Ravn et al., 2012), SEN2006/DENNIS2013 (Sen, 2006; paper unavailable, so data extracted from Dennis & Dowswell, 2013), SMALL2000/2006 (one study reported across two papers: Small et al., 2000; Small et al., 2006), SPITTLE2010/2009/SPENCERSMITH2012 (one study reported across three papers: Spittle et al., 2009; Spittle et al., 2010; Spencer-Smith et al., 2012), STAMP1995 (Stamp et al., 1995), WEBSTER2003 (Webster et al., 2003). All of these studies were published in peer-reviewed journals between 1994 and 2013. In addition, 33 studies were excluded from the review. The most common reasons for exclusion were that data could not be extracted (for instance, because means and standard deviations were not reported), or there were no mental health outcomes reported, or the studies were not RCTs. Further information about both included and excluded studies can be found in Appendix 18.

For the review of protocols for women following stillbirth, four cohort studies reported across six papers met the eligibility criteria for this review: CACCIATORE2008 (Cacciatore et al., 2008), GRAVENSTEEN2013 (Gravensteen et al., 2013), HUGHES2002/TURTON2009 (Hughes et al., 2002; Turton et al., 2009), RADESTAD2009A/SURKAN2008 (Rådestad et al., 2009a; Surkan et al., 2008). All of these studies were published in peer-reviewed journals between 2002 and 2013. In addition, two studies were excluded (CRAWLEY2013 [Crawley et al., 2013], RADESTAD2009B [Rådestad et al., 2009b]) as data could not be extracted as there was not a sufficient comparison group (>90% saw and held the stillborn infant). Further information about both included and excluded studies can be found in Appendix 18.

Of the 22 included RCTs, there was one study (N=228) involving a comparison of post-miscarriage self-help and treatment as usual (Table 32). The term post-

[11]Here and elsewhere in the guideline, each study considered for review is referred to by a study ID in capital letters (primary author and date of study publication, except where a study is in press or only submitted for publication, then a date is not used).

miscarriage is used as a proxy for loss of baby during pregnancy due to miscarriage, termination due to fetal abnormality, or stillbirth.

There was one study (N=117) that compared social support (peer-mediated support) with treatment as usual (Table 33). This study did not clarify risk factors but defined the sample as 'at risk'.

There were three studies (N=360) that involved a comparison between psychologically (CBT/IPT)-informed psychoeducation and treatment as usual or enhanced treatment as usual for women with psychosocial risk factors, for teenage mothers, or for women classified as 'at risk' but where risk factors were not defined. Two studies (N=1,140) compared a psychoeducational booklet and treatment as usual or enhanced treatment as usual for women with psychosocial risk factors. Four studies (N=844) compared non-mental health-focused education and support and treatment as usual or enhanced treatment as usual for women with a range of risk factors including psychosocial risk factors, preterm delivery and low birthweight baby, and multiple (twin) pregnancy. Five studies (N=1,146) involved a comparison of home visits and treatment as usual predominantly for women with psychosocial risk factors, but also including teenage mothers and one study which examined women at risk of mental health problems due to preterm delivery. One study (N=1,041) compared post-delivery discussion and enhanced treatment as usual (Table 34) for women who had had an operative delivery.

Four studies (N=799) compared mother–infant relationship interventions and treatment as usual (Table 35) for women with psychosocial risk factors or with premature or low birthweight babies.

There was one study (N=34) that involved a comparison between case management and individualized treatment and treatment as usual (Table 36) for women who had preterm delivery and low birthweight babies.

Four studies (N=2,772) compared mental health outcomes in women who saw and/or held their stillborn infants compared with those who did not (Table 37).

Table 32: Study information table for trials included in the prevention (risk factors identified) meta-analysis of self-help versus any alternative management strategy

	Post-miscarriage self-help versus TAU
Total number of trials (number of participants)	1 (228)
Study ID	KERSTING2013
Country	European German-speaking countries
Mean age of participants (years)	34.2

Risk factor(s)	Miscarriage, termination due to fetal abnormality, or stillbirth
Timing of intervention	Post-miscarriage
Mode of delivery	Internet
Format	Individual
Intensity (number of sessions)	Low (0 sessions of contact with professional; 5 internet sessions [10 essays])
Length of intervention (weeks)	5
Time points	Post-treatment
Setting	Internet
Intervention	Internet-based CBT-informed self-help
Comparison	Waitlist
Note. Abbreviations: TAU=Treatment as usual	

Table 33: Study information table for trials included in the prevention (risk factors identified) meta-analysis of social support versus any alternative management strategy

	Social support versus TAU
Total number of trials (number of participants)	1 (117)
Study ID	HARRIS2006/DENNIS2013
Country	UK
Mean age of participants (years)	NR
Risk factor(s)	Unclear ('at-risk')
Timing of intervention	Antenatal and postnatal
Mode of delivery	Face-to-face
Format	Individual and group
Intensity (number of sessions)	NR
Length of intervention (weeks)	NR
Time points	Post-treatment
Setting	NR
Intervention	Peer-mediated support (including one-to-one befriending and psychoeducational group meetings)
Comparison	TAU
Note. Abbreviations: NR=Not reported; TAU=Treatment as usual	

Table 34: Study information table for trials included in the prevention (risk factors identified) meta-analysis of education or support versus any alternative management strategy

	Psychologically (CBT/ IPT)-informed psychoeducation versus TAU or Enhanced TAU	Psychoeducational booklet versus TAU or Enhanced TAU	Non-mental health-focused education and support versus TAU or Enhanced TAU	Home visits versus TAU	Post-delivery discussion versus Enhanced TAU
Total number of trials (number of participants)	3 (360)	2 (1,140)	4 (844)	5 (1,146)	1 (1,041)
Study ID	(1) BRUGHA2000 (2) GORMAN1997/ DENNIS2013 (3) PHIPPS2013	(1) HOWELL2012 (2) WEBSTER2003	(1) KIEFFER2013 (2) MELNYK2006 (3) SEN2006/ DENNIS2013 (4) STAMP1995	(1) ARACENA2009 (2) BARLOW2007 (3) BARNET2007 (4) EASTERBROOKS2013 (5) SPITTLE2010/2009/ SPENCERSMITH2012	SMALL2000/2006
Country	(1) UK (2)-(3) US	(1) US (2) Australia	(1)-(2) US (3) UK (4) Australia	(1) Chile (2) UK (3)-(4) US (5) Australia	Australia
Mean age of participants (years)	(1) Median: 19 (2) NR (3) Median: 16	(1) 28 (2) 27.2	(1) NR (2) 27.8 (3) NR (4) 26.5	(1) 17.2 (2) NR (3) 16.9 (4) 18.7 (5) NR	NR
Risk factor(s)	(1) Psychosocial (2) Unclear ('at-risk') (3) Adolescence and psychosocial	(1) Psychosocial (2) Psychosocial and (family) history of mental health problems	(1) Psychosocial (2) Preterm delivery and low birthweight (3) Multiple (twin) pregnancy (4) Uncertain ('at risk')	(1) Adolescence and psychosocial (2) Psychosocial and (family) history of mental health problems (3)-(4) Adolescence and psychosocial (5) Preterm delivery	Operative delivery
Timing of intervention	(1) Antenatal	(1) Postnatal (2) Antenatal	(1) Antenatal and postnatal	(1)-(3) Antenatal and postnatal	Postnatal

	(2) Antenatal and postnatal (3) Antenatal		(2) Postnatal (3)-(4) Antenatal and postnatal	(4) Antenatal (5) Postnatal	
Mode of delivery	(1)-(3) Face-to-face	(1) Booklet and telephone (2) Booklet	(1) Face-to-face (2) Written and audiotaped (3)-(4) Face-to-face	(1)-(5) Face-to-face	Face-to-face
Format	(1) Group (2) Individual (3)Individual and group	(1)-(2) Individual	(1) Individual and group (2) Individual (3) Individual and group (4) Group	(1)-(5) Individual	Individual
Intensity (number of sessions)	(1)-(3) Low (5-6 sessions)	(1)-(2) Low (1-2 sessions)	(1) Moderate (11 sessions) (2) Low (0 sessions contact with healthcare professional; 4 sessions of written and audiotaped information) (3)-(4) Moderate (8-10 sessions)	(1) Moderate (12 sessions) (2)-(3) High (41-45 sessions) (4) NR (5) Moderate (9 sessions)	Low (single session)
Length of intervention (weeks)	(1) 6 (2) NR (3) 5	(1) 2 (2) NR	(1) 17 (2)-(3) NR (4) 13	(1) NR (2) 78 (3) 117 (4) NR (5) 52	Single session
Time points	(1) Post-treatment (2) Post-treatment; Intermediate follow-up (3) Post-treatment	(1) Post-treatment; Short follow-up; Intermediate follow-up (2) Post-treatment	(1) Post-treatment (2) Post-treatment; Mid-treatment (3) Post-treatment; Short follow-up; Intermediate follow-up; Long follow-up (4) Post-treatment	(1)-(3) Post-treatment (4) First measurement (5) First measurement; Very long follow-up	First measurement; Very long follow-up
Setting	(1) Hospital (2)-(3) NR	(1) Hospital and telephone (2) Hospital	(1) Community and home (2) Hospital	(1)-(5) Home	Hospital

			(3) Home, hospital and clinic (secondary) (4) Clinic (primary)		
Intervention	(1) CBT-informed psychoeducation (2)-(3) IPT-informed psychoeducation	(1) Psychoeducational booklet and telephone support (2) Psychoeducational booklet	(1) Non-mental health-focused education and support group and home visits (2) Non-mental health-focused education and support (booklet and audiotaped) (3) Non-mental health-focused education and support group and home visits (4) Non-mental health-focused education and support group	(1)-(5) Home visits	Midwife-led post-delivery discussion
Comparison	(1)-(2) TAU (3) Enhanced TAU (non-mental health-focused education and support [booklet])	(1) Enhanced TAU (non-mental health-focused education and support [booklet]) (2) TAU	(1) Enhanced TAU (non-mental health-focused education and support without the focus on healthy eating and exercise) (2) Enhanced TAU (non-mental health-focused information) (3)-(4) TAU	(1)-(5) TAU	Enhanced TAU (Non-mental health-focused information [booklet])
Note. Abbreviations: NR=Not reported; TAU=Treatment as usual					

Table 35: Study information table for trials included in the prevention (risk factors identified) meta-analysis of mother–infant relationship interventions versus any alternative management strategy

	Mother-infant relationship interventions versus TAU
Total number of trials (number of participants)	4 (799)
Study ID	(1) COOPER2009 (2) MEIJSSEN2010A/2010B/2011 (3) NEWNHAM2009 (4) RAVN2012
Country	(1) South Africa (2) Netherlands (3) Australia (4) Norway
Mean age of participants (years)	(1) 25.9 (2) 32.2 (3) 31.5 (4) 30.9
Risk factor(s)	(1) Psychosocial (2)-(4) Preterm delivery and/or low birthweight
Timing of intervention	(1) Antenatal and postnatal (2)-(4) Postnatal
Mode of delivery	(1)-(4) Face-to-face
Format	(1)-(4) Individual
Intensity (number of sessions)	(1) High (16 sessions) (2)-(4) Moderate (8-11 sessions)
Length of intervention (weeks)	(1)-(2) NR (3) 15 (4) 14
Time points	(1) Post-treatment; First measurement; Long follow-up (2) First measurement; Long follow-up (3) Post-treatment; Short follow-up (4) First measurement; Long follow-up
Setting	(1)-(2) Home (3)-(4) Hospital and home
Intervention	(1)-(4) Mother-infant relationship interventions
Comparison	(1)-(4) TAU
Note. Abbreviations: NR=Not reported; TAU=Treatment as usual	

Table 36: Study information table for trials included in the prevention (risk factors identified) meta-analysis of other psychosocial interventions versus any alternative management strategy

	Case management and individualized treatment versus TAU
Total number of trials (number of participants)	1 (34)
Study ID	MEYER1994
Country	US
Mean age of participants (years)	27.9
Risk factor(s)	Preterm delivery and low birthweight
Timing of intervention	Postnatal
Mode of delivery	Face-to-face
Format	Individual
Intensity (number of sessions)	Moderate (median: 10 sessions)
Length of intervention (weeks)	Median: 5
Time points	Post-treatment
Setting	Hospital
Intervention	Case management and individualized treatment
Comparison	TAU
Note. Abbreviations: NR=Not reported; TAU=Treatment as usual	

Table 37: Study information table for trials included in the prevention (risk factors identified) meta-analysis of protocols following stillbirth

	Seeing and/or holding stillborn infant versus not seeing or not holding stillborn infant
Total number of trials (number of participants)	4 (2,772)
Study ID	(1) CACCIATORE2008 (2) GRAVENSTEEN2013 (3) HUGHES2002/TURTON2009 (4) RADESTAD2009A/SURKAN2008
Country	(1) US (72%); UK (11%); Australia (9%); Canada (5%) (2) Norway (3) UK (4) Sweden
Study design	(1)-(2) and (4) Cohort (retrospective) (3) Nested cohort within case-control
Recruitment approach	(1) SR of internet search engines and directories to identify organisations to recruit women affected by stillbirth to respond to an online questionnaire (2) Hospital records used to identify verified diagnosis of stillbirth from 1 January 1990 to 31 December 2003 and a postal invitation sent to potential participants

	(3) Women who had previously experienced a stillbirth who were pregnant with another child and attended an antenatal clinic at one of three district general hospitals. (4) Swedish population-based Medical Birth Register was used to identify all women who had had a stillborn baby in Sweden in 1991
Timing (length of time since stillbirth)	(1) 51% ≤1 year; 15% 1-2 years; 9% 2-3 years; 25% ≥3 years (2) 5-18 years after stillbirth (mean: 10.8 years) (3) Unclear (51% conceived less than 12 months after loss and 49% more than 12 months after loss) (4) 3 years after the stillbirth
Pregnancy status at time of participation	(1) 286 women (12%) pregnant (2) None of the women were pregnant at follow-up; mean of 2.2 live-born children (3) All of the women were pregnant at time of study (4) NR
Mean gestational age at time of stillbirth	(1) NR (inclusion criteria >20) (2) NR (inclusion criteria ≥23) (3) NR (inclusion criteria >18) (4) NR (inclusion criteria >28 weeks. 39% 28-37 weeks; 50% 38-42 weeks; 10% >42 weeks)
Note. Abbreviations: NR=Not reported; SR=Systematic review	

7.4.4 Clinical evidence for preventative effects on depression outcomes for women with identified risk factors (by intervention)

Summary of findings can be found in the tables presented in this section. The full GRADE evidence profiles and associated forest plots can be found in Appendix 22 and Appendix 19, respectively.

Depression: post-miscarriage self-help versus treatment as usual

There was single study (N=228) evidence for a moderate preventative benefit of post-miscarriage self-help on depression mean symptoms ($p < 0.00001$). However, the confidence in this effect estimate is low due to risk of bias (statistically significant group differences at baseline) and imprecision (optimal information size [N=400] is not met). The outcome measure is also a subscale of a global severity measure (Brief Symptom Inventory [BSI]: Depression) rather than a depression-specific scale (Table 38).

Table 38: Summary of findings table for effects of post-miscarriage self-help compared with treatment as usual on preventing depression outcomes in women with identified risk factors

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Depression: post-miscarriage self-help versus TAU				
Depression mean symptoms Post-treatment – ITT analysis (at-risk populations) BSI: Depression Follow-up: mean 5 weeks		The mean depression mean symptoms post-treatment – ITT analysis (at-risk populations) in the intervention groups was 0.64 standard deviations lower (0.91 to 0.37 lower)		228 (1 study)	⊕⊕⊖⊖ low[1,2]	SMD -0.64 (-0.91 to -0.37)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Risk of bias due to statistically significant group differences at baseline
[2] Total population size is less than 400 (a threshold rule-of-thumb)

Depression: social support versus treatment as usual

There was very low quality, single study (N=65) evidence for a large preventative benefit of social support on depression diagnosis (p=0.01) in women at risk of developing postnatal depression, when using an available case analysis approach. However, ITT analysis of this outcome measure revealed no evidence for statistically or clinically significant effects of social support on depression diagnosis (p=0.22). Moreover, there are risk of bias concerns with this study due to non-blind outcome assessment (Table 39).

Table 39: Summary of findings table for effects of social support compared with treatment as usual on preventing depression outcomes in women with identified risk factors

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Depression: Social support versus TAU				
Depression diagnosis post-treatment – ITT analysis (at-risk populations) Schedules for Clinical Assessment in Neuropsychiatry (SCAN) Follow-up: mean 12 weeks	Study population		RR 0.85 (0.65 to 1.1)	117 (1 study)	⊕⊖⊖⊖ very low1,2,3	
	714 per 1000	607 per 1000 (464 to 786)				
	Moderate					
	714 per 1000	607 per 1000 (464 to 785)				
Depression diagnosis post-treatment – available case analysis (at-risk populations) SCAN Follow-up: mean 12 weeks	Study population		RR 0.37 (0.17 to 0.8)	65 (1 study)	⊕⊖⊖⊖ very low1,2	
	543 per 1000	**201 per 1000** (92 to 434)				
	Moderate					
	543 per 1000	**201 per 1000** (92 to 434)				

*The basis for the **assumed risk** (for example, the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% CI) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Risk of bias due to non-blind outcome assessment
[2] Total number of events is less than 300 (a threshold rule-of-thumb)
[3] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Depression: Psychologically (CBT/IPT)-informed psychoeducation versus treatment as usual or enhanced treatment as usual

The evidence for psychologically (CBT/IPT)-informed psychoeducation as a preventative intervention for women at-risk of developing postnatal depression was inconsistent (Table 40). There was evidence from three studies (N=320-360) for moderate to large effects of psychoeducation on preventing depression diagnosis (using either ITT [p=0.08] or available case [p=0.05] data analysis). However, the confidence in this effect estimate is low due to very serious imprecision (small event

rate and the 95% CI included both no effect and appreciable benefit). This effect was also not maintained at intermediate (17-24 weeks post-intervention) follow-up (p=0.51-0.53). In addition, no clinically or statistically significant preventative effects were observed on depression symptomatology at endpoint (p=0.41-0.66) or intermediate follow-up (p=0.63-1), or depression mean symptoms at endpoint (p=0.86) or intermediate follow-up (p=0.96).

Table 40: Summary of findings table for effects of psychologically (CBT/IPT)-informed psychoeducation compared with treatment as usual or enhanced treatment as usual on preventing depression outcomes in women with identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Depression: Psychologically (CBT/IPT)-informed psychoeducation versus TAU or Enhanced TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Depression diagnosis post-treatment – ITT analysis (at-risk populations) SCAN, SCID or Structured Clinical Interview for Childhood Diagnoses (KID-SCID) Follow-up: mean 27 weeks	Study population 229 per 1000	158 per 1000 (103 to 241)	RR 0.69 (0.45 to 1.05)	360 (3 studies)	⊕⊕⊖⊖ low1,2	
	Moderate 333 per 1000	230 per 1000 (150 to 350)				
Depression diagnosis post-treatment – available case analysis (at-risk populations) SCAN, SCID or KID-SCID Follow-up: mean 27 weeks	Study population 132 per 1000	63 per 1000 (30 to 133)	RR 0.48 (0.23 to 1.01)	320 (3 studies)	⊕⊕⊖⊖ low1,2	
	Moderate 227 per 1000	109 per 1000 (52 to 229)				
Depression symptomatology Post-treatment – ITT analysis (at-risk populations) EPDS ≥11/12	Study population 299 per 1000	254 per 1000 (174 to 374)	RR 0.85 (0.58 to 1.25)	254 (2 studies)	⊕⊕⊖⊖ low1,2	
	Moderate 370 per 1000	315 per 1000 (215 to 462)				

Follow-up: mean 27 weeks						
Depression symptomatology Post-treatment – available case analysis (at-risk populations) EPDS ≥1/12 Follow-up: mean 27 weeks	Study population		RR 0.88 (0.49 to 1.57)	221 (2 studies)	⊕⊕⊖⊖ low1,2	
	183 per 1000	161 per 1000 (90 to 288)				
	Moderate					
	171 per 1000	150 per 1000 (84 to 268)				
Depression mean scores post-treatment – available case analysis (at-risk populations) EPDS		The mean depression mean scores post-treatment – available case analysis (at-risk populations) in the intervention groups was 0.06 standard deviations lower (0.75 lower to 0.62 higher)		33 (1 study)	⊕⊕⊖⊖ low1	SMD -0.06 (-0.75 to 0.62)
Depression diagnosis intermediate follow-up (17-24 weeks post-intervention) – ITT analysis (at-risk populations) SCID Follow-up: mean 20 weeks	Study population		RR 0.77 (0.33 to 1.75)	45 (1 study)	⊕⊕⊖⊖ low1,2	
	381 per 1000	293 per 1000 (126 to 667)				
	Moderate					
	381 per 1000	293 per 1000 (126 to 667)				
Depression diagnosis intermediate follow-up (17-24 weeks post-intervention) – available case analysis (at-risk populations) SCID Follow-up: mean 20 weeks	Study population		RR 0.64 (0.17 to 2.46)	37 (1 study)	⊕⊕⊖⊖ low1,2	
	235 per 1000	151 per 1000 (40 to 579)				
	Moderate					
	235 per 1000	150 per 1000 (40 to 578)				
Depression symptomatology Intermediate follow-up (17-24 weeks post-intervention) – ITT analysis (at-risk populations) EPDS >12 Follow-up: mean 20 weeks	Study population		RR 1.17 (0.62 to 2.2)	45 (1 study)	⊕⊕⊖⊖ low1,2	
	429 per 1000	501 per 1000 (266 to 943)				
	Moderate					
	429 per 1000	502 per 1000 (266 to 944)				
	Study population					

Depression symptomatology Intermediate follow-up (17-24 weeks post-intervention) – available case analysis (at-risk populations) EPDS >12 Follow-up: mean 20 weeks	200 per 1000 Moderate 200 per 1000	200 per 1000 (48 to 836) 200 per 1000 (48 to 836)	RR 1 (0.24 to 4.18)	30 (1 study)	⊕⊕⊖⊖ low1,2	
Depression mean scores Intermediate follow-up (17-24 weeks post-intervention) – available case analysis (at-risk populations) EPDS Follow-up: mean 20 weeks		The mean depression mean scores intermediate follow-up (17-24 weeks post-intervention) – available case analysis (at-risk populations) in the intervention groups was 0.02 standard deviations lower (0.74 lower to 0.7 higher)		30 (1 study)	⊕⊕⊖⊖ low1,2	SMD -0.02 (-0.74 to 0.7)

*The basis for the **assumed risk** (for example, the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% CI) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total number of events is less than 300 (a threshold rule-of-thumb)
[2] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Depression: Psychoeducational booklet versus treatment as usual or enhanced treatment as usual

There was low to very low quality evidence from up to two studies (N=1,140) for moderate effects of a psychoeducational booklet on preventing depression symptomatology (p=0.10-0.11) in women with psychosocial risk factors when an available case analysis approach was used (Table 41). However, moderate to low quality evidence from ITT analyses provided no evidence for psychoeducation as an intervention to prevent depression symptomatology (p=0.12-0.46).

Table 41: Summary of findings table for effects of psychoeducational booklet compared with treatment as usual or enhanced treatment as usual on preventing depression outcomes in women with identified risk factors

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk Control	Corresponding risk Depression: Psychoeducational booklet versus TAU or Enhanced TAU				
Depression symptomatology Post-treatment – ITT analysis (at-risk populations) EPDS ≥10/12 Follow-up: mean 3 weeks	Study population 419 per 1000	377 per 1000 (331 to 431)	RR 0.9 (0.79 to 1.03)	1,140 (2 studies)	⊕⊕⊕⊖ moderate[1]	
	Moderate 409 per 1000	368 per 1000 (323 to 421)				
Depression symptomatology Post-treatment – available case analysis (at-risk populations) EPDS ≥10/12 Follow-up: mean 3 weeks	Study population 208 per 1000	152 per 1000 (106 to 220)	RR 0.73 (0.51 to 1.06)	838 (2 studies)	⊕⊖⊖⊖ very low[1,2,3]	
	Moderate 218 per 1000	159 per 1000 (111 to 231)				
Depression symptomatology Short Follow-up (9-16 weeks post-intervention) – ITT analysis (at-risk populations) EPDS ≥10 Follow-up: mean 13 weeks	Study population 222 per 1000	196 per 1000 (142 to 273)	RR 0.88 (0.64 to 1.23)	540 (1 study)	⊕⊕⊖⊖ low[1,2]	
	Moderate 222 per 1000	195 per 1000 (142 to 273)				
Depression symptomatology Short Follow-up (9-16 weeks post-intervention) – available case analysis (at-risk populations) EPDS ≥10 Follow-up: mean 13 weeks	Study population 132 per 1000	85 per 1000 (50 to 143)	RR 0.64 (0.38 to 1.08)	479 (1 study)	⊕⊕⊖⊖ low[2,3]	
	Moderate 132 per 1000	84 per 1000 (50 to 143)				
Depression symptomatology Intermediate follow-up (17-24 weeks post-intervention) – ITT analysis (at-risk populations) EPDS ≥10 Follow-up: mean 26 weeks	Study population 333 per 1000	277 per 1000 (217 to 360)	RR 0.83 (0.65 to 1.08)	540 (1 study)	⊕⊕⊖⊖ low[2,3]	
	Moderate 333 per 1000	276 per 1000 (216 to 360)				
	Study population					

Depression symptomatology Intermediate follow-up (17-24 weeks post-intervention) – available case analysis (at-risk populations) EPDS ≥10 Follow-up: mean 26 weeks	139 per 1000 Moderate 139 per 1000	89 per 1000 (51 to 153) 89 per 1000 (51 to 153)	RR 0.64 (0.37 to 1.1)	423 (1 study)	⊕⊕⊖⊖ low[2,3]

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Risk of bias due to statistically significant group differences at baseline
[2] Total number of events is less than 300 (a threshold rule-of-thumb)
[3] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Depression: non-mental health-focused education and support versus treatment as usual or enhanced treatment as usual

Low quality evidence from up to two studies (N=306) suggests that non-mental health-focused education and support may be more effective than treatment as usual or enhanced treatment as usual at preventing depression symptomatology for women with multiple births or at risk of developing postnatal depression (no further details reported) with moderate effects observed at endpoint (p=0.07-0.15) and moderate to large effects observed at short-term (9-16 weeks post-intervention) follow-up (p=0.09). However, effects were not maintained at intermediate (p=0.77-0.81) or long-term (p=0.40-0.72) follow-ups, and there was no evidence for statistically or clinically significant preventative benefits for depression mean symptoms at any time point (p=0.09-0.64) (Table 42).

Table 42: Summary of findings table for effects of non-mental health-focused education and support compared with treatment as usual or enhanced treatment as usual on preventing depression outcomes in women with identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk / Corresponding risk Control	Depression: Non-mental health-focused education and support versus TAU or Enhanced TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Depression symptomatology Post-treatment – ITT analysis (at-risk populations) EPDS >12 Follow-up: 6-13 weeks	Study population 320 per 1000	224 per 1000 (141 to 365)	RR 0.7 (0.44 to 1.14)	306 (2 studies)	⊕⊕⊖⊖ low[1,2]	
	Moderate 316 per 1000	221 per 1000 (139 to 360)				
Depression symptomatology Post-treatment – available case analysis (at-risk populations) EPDS >12 Follow-up: 6-13 weeks	Study population 188 per 1000	107 per 1000 (58 to 197)	RR 0.57 (0.31 to 1.05)	261 (2 studies)	⊕⊕⊖⊖ low[1,2]	
	Moderate 188 per 1000	107 per 1000 (58 to 197)				
Depression mean scores post-treatment – ITT analysis (at-risk populations) CES-D Follow-up: mean 28 weeks		The mean depression mean scores post-treatment – ITT analysis (at-risk populations) in the intervention groups was 0.13 standard deviations lower (0.37 lower to 0.1 higher)		275 (1 study)	⊕⊕⊖⊖ low[3,4]	SMD -0.13 (-0.37 to 0.1)
Depression mean scores post-treatment – available case analysis (at-risk populations) BDI or EPDS		The mean depression mean scores post-treatment – available case analysis (at-risk populations) in the intervention groups was 0.14 standard deviations lower		370 (2 studies)	⊕⊕⊕⊖ moderate[3]	SMD -0.14 (-0.34 to 0.07)

		(0.34 lower to 0.07 higher)				
Depression symptomatology Short Follow-up (9-16 weeks post-intervention) – ITT analysis (at-risk populations) EPDS >12 Follow-up: mean 6 weeks	Study population 402 per 1000 Moderate 402 per 1000	274 per 1000 (177 to 427) 273 per 1000 (177 to 426)	RR 0.68 (0.44 to 1.06)	162 (1 study)	⊕⊕⊖⊖ low[1,2]	
Depression symptomatology Short Follow-up (9-16 weeks post-intervention) – available case analysis (at-risk populations) – Non-mental health-focused education and support EPDS >12 Follow-up: mean 12 weeks	Study population 222 per 1000 Moderate 222 per 1000	107 per 1000 (47 to 249) 107 per 1000 (47 to 249)	RR 0.48 (0.21 to 1.12)	128 (1 study)	⊕⊕⊖⊖ low[1,2]	
Depression mean scores Short Follow-up (9-16 weeks post-intervention) – available case analysis (at-risk populations) EPDS Follow-up: mean 12 weeks		The mean depression mean scores short follow-up (9-16 weeks post-intervention) – available case analysis (at-risk populations) in the intervention groups was 0.21 standard deviations lower (0.56 lower to 0.13 higher)		128 (1 study)	⊕⊕⊖⊖ low[2,3]	SMD -0.21 (-0.56 to 0.13)
Depression symptomatology Intermediate follow-up (17-24 weeks post-intervention) – ITT analysis (at-risk populations) EPDS >12 Follow-up: 20-24 weeks	Study population 294 per 1000 Moderate 290 per 1000	268 per 1000 (129 to 556) 264 per 1000 (128 to 548)	RR 0.91 (0.44 to 1.89)	306 (2 studies)	⊕⊖⊖⊖ very low[1,2,5]	
Depression symptomatology Intermediate follow-up (17-24 weeks post-intervention) – available case analysis (at-risk populations) EPDS >12 Follow-up: 20-24 weeks	Study population 143 per 1000 Moderate 142 per 1000	120 per 1000 (39 to 376) 119 per 1000 (38 to 373)	RR 0.84 (0.27 to 2.63)	254 (2 studies)	⊕⊖⊖⊖ very low[1,2,5]	

Depression mean scores Intermediate follow-up (17-24 weeks post-intervention) – available case analysis (at-risk populations) EPDS Follow-up: mean 24 weeks		The mean depression mean scores intermediate follow-up (17-24 weeks post-intervention) – available case analysis (at-risk populations) in the intervention groups was 0.3 standard deviations lower (0.64 lower to 0.04 higher)		133 (1 study)	⊕⊕⊖⊖ low[2,3]	SMD -0.3 (-0.64 to 0.04)
Depression symptomatology Long Follow-up (25-103 weeks post-intervention) – ITT analysis (at-risk populations) EPDS >12 Follow-up: mean 52 weeks	Study population		RR 0.84 (0.57 to 1.25)	162 (1 study)	⊕⊕⊖⊖ low[1,2]	
	415 per 1000	348 per 1000 (236 to 518)				
	Moderate					
	415 per 1000	349 per 1000 (237 to 519)				
Depression symptomatology Long Follow-up (25-103 weeks post-intervention) – available case analysis (at-risk populations) EPDS >12 Follow-up: mean 52 weeks	Study population		RR 0.87 (0.42 to 1.83)	123 (1 study)	⊕⊕⊖⊖ low[1,2]	
	200 per 1000	174 per 1000 (84 to 366)				
	Moderate					
	200 per 1000	174 per 1000 (84 to 366)				
Depression mean scores Long Follow-up (25-103 weeks post-intervention) -Available case analysis (at-risk populations) EPDS Follow-up: mean 52 weeks		The mean depression mean scores long follow-up (25-103 weeks post-intervention) - available case analysis (at-risk populations) in the intervention groups was 0.08 standard deviations lower (0.44 lower to 0.27 higher)		123 (1 study)	⊕⊕⊖⊖ low[3]	SMD -0.08 (-0.44 to 0.27)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.

Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total number of events is less than 300 (a threshold rule-of-thumb)
[2] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
[3] Total population size is less than 400 (a threshold rule-of-thumb)
[4] Paper omits data
[5] There is evidence of substantial heterogeneity of study effect sizes

Depression: home visits versus treatment as usual

Using an available case data analysis approach there is single study (N=77) evidence suggesting that home visits may be more effective than treatment as usual at preventing depression symptomatology at very long (>104 weeks post-intervention) follow-up (p=0.28). However, confidence in this effect estimate is very low due to risk of bias concerns (statistically significant group differences in depression symptomatology at baseline) and very serious imprecision (optimal information size [that is, 300 events] is not met and 95% CI includes no effect, appreciable benefit and appreciable harm). Moreover, the ITT analysis of this outcome measure is not statistically or clinically significant (p=0.60) and there is no evidence (from up to three studies; N=684) for statistically or clinically significant effects on depression symptomatology at endpoint or first measurement (p=0.42-0.87) or depression mean symptoms at very long follow-up (p=0.11), or for clinically significant effects on mean depression symptoms at endpoint (p=0.04) (Table 43).

Table 43: Summary of findings table for effects of home visits compared with treatment as usual on preventing depression outcomes in women with identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Depression: Home visits versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Depression symptomatology Post-treatment – ITT analysis (at-risk populations) CES-D ≥21 or HADS – Depression >7 Follow-up: 52-117 weeks	Study population 434 per 1000	408 per 1000 (195 to 851)	RR 0.94 (0.45 to 1.96)	204 (2 studies)	⊕⊖⊖⊖ very low[1,2,3,4,5]	
	Moderate 429 per 1000	403 per 1000 (193 to 841)				
Depression symptomatology Post-treatment – available case analysis (at-risk populations) CES-D ≥16/21 or HADS – Depression >7 Follow-up: 52-117 weeks	Study population 332 per 1000	259 per 1000 (146 to 468)	RR 0.78 (0.44 to 1.41)	684 (3 studies)	⊕⊖⊖⊖ very low[1,3,4,6]	
	Moderate 256 per 1000	200 per 1000 (113 to 361)				
Depression mean scores post-treatment – available case analysis (at-risk populations) CES-D or HADS – Depression Follow-up: mean 52 weeks		The mean depression mean scores post-treatment – available case analysis (at-risk populations) in the intervention groups was 0.38 standard deviations lower (0.75 to 0.01 lower)		621 (2 studies)	⊕⊖⊖⊖ very low[1,7]	SMD -0.38 (-0.75 to -0.01)
Depression symptomatology Very long Follow-up (>104 weeks post-intervention) – ITT analysis (at-risk populations) HADS – Depression ≥8 Follow-up: mean 104 weeks	Study population 458 per 1000	412 per 1000 (270 to 618)	RR 0.90 (0.59 to 1.35)	120 (1 study)	⊕⊖⊖⊖ very low[1,3,4,5]	
	Moderate 158 per 1000	142 per 1000 (93 to 213)				

Depression symptomatology Very long Follow-up (>104 weeks post-intervention) – available case analysis (at-risk populations) HADS – Depression ≥8 Follow-up: mean 104 weeks	Study population		RR 0.49 (0.13 to 1.81)	77 (1 study)	⊕⊖⊖⊖ very low[1,3,4,5]	
	158 per 1000	77 per 1000 (21 to 286)				
	Moderate					
	158 per 1000	77 per 1000 (21 to 286)				
Depression mean scores Very long Follow-up (>104 weeks post-intervention) – available case analysis (at-risk populations) HADS – Depression Follow-up: mean 104 weeks		The mean depression mean scores very long follow-up (>104 weeks post-intervention) – available case analysis (at-risk populations) in the intervention groups was 0.37 standard deviations lower (0.82 lower to 0.08 higher)		77 (1 study)	⊕⊖⊖⊖ very low[1,4,5,8]	SMD -0.37 (-0.82 to 0.08)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Risk of bias due to statistically significant group differences at baseline
[2] There is evidence of considerable heterogeneity of study effect sizes
[3] Total number of events is less than 300 (a threshold rule-of-thumb)
[4] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
[5] Paper omits data
[6] There is evidence of moderate heterogeneity of study effect sizes
[7] There is evidence of substantial heterogeneity of study effect sizes
[8] Total population size is less than 400 (a threshold rule-of-thumb)

Depression: post-delivery discussion versus enhanced treatment as usual

There was no evidence (Table 44) that a post-delivery discussion was more effective than enhanced treatment as usual (non-mental health-focused information [booklet]) at preventing depression in women following an operative delivery (p=0.23-0.87).

Table 44: Summary of findings table for effects of post-delivery discussion compared with enhanced treatment as usual on preventing depression outcomes in women with identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk: Control	Corresponding risk: Depression: post-delivery discussion versus Enhanced TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Depression symptomatology Post-treatment – ITT analysis (at-risk populations) EPDS ≥13 Follow-up: mean 26 weeks	Study population 263 per 1000	258 per 1000 (210 to 316)	RR 0.98 (0.8 to 1.2)	1,041 (1 study)	⊕⊕⊕⊖ moderate[1]	
	Moderate 263 per 1000	258 per 1000 (210 to 316)				
Depression symptomatology Post-treatment – available case analysis (at-risk populations) EPDS ≥13 Follow-up: mean 26 weeks	Study population 145 per 1000	174 per 1000 (129 to 235)	RR 1.2 (0.89 to 1.62)	916 (1 study)	⊕⊕⊖⊖ low[1,2]	
	Moderate 145 per 1000	174 per 1000 (129 to 235)				
Depression mean scores post-treatment – available case analysis (at-risk populations) EPDS Follow-up: mean 26 weeks		The mean depression mean scores post-treatment – available case analysis (at-risk populations) in the intervention groups was 0.08 standard deviations higher (0.05 lower to 0.21 higher)		916 (1 study)	⊕⊕⊕⊕ high	SMD 0.08 (-0.05 to 0.21)
Depression symptomatology Very long Follow-up (>104 weeks post-intervention) – ITT analysis (at-risk populations) EPDS ≥13 Follow-up: 208-312 weeks	Study population 568 per 1000	574 per 1000 (517 to 636)	RR 1.01 (0.91 to 1.12)	1,041 (1 study)	⊕⊕⊕⊕ high	
	Moderate 568 per 1000	574 per 1000 (517 to 636)				
	Study population					

Depression symptomatology Very long Follow-up (>104 weeks post-intervention) – available case analysis (at-risk populations) EPDS ≥13 Follow-up: 208-312 weeks	167 per 1000	158 per 1000 (108 to 233)				
	Moderate					
	167 per 1000	159 per 1000 (109 to 234)	RR 0.95 (0.65 to 1.4)	534 (1 study)	⊕⊕⊖⊖ low[1,2]	
Depression mean scores Very long Follow-up (>104 weeks post-intervention) – available case analysis (at-risk populations) EPDS Follow-up: 208-312 weeks		The mean depression mean scores very long follow-up (>104 weeks post-intervention) – available case analysis (at-risk populations) in the intervention groups was 0.08 standard deviations lower (0.25 lower to 0.09 higher)		534 (1 study)	⊕⊕⊕⊕ high	SMD -0.08 (-0.25 to 0.09)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total number of events is less than 300 (a threshold rule-of-thumb)
[2] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Depression: mother-infant relationship interventions versus treatment as usual

The evidence for mother–infant relationship interventions preventing depression in women with psychosocial risk factors or who had a preterm delivery and/or low birthweight baby was very inconsistent (Table 45). There was single study (N=106) evidence for large harms associated with mother–infant relationship interventions for women who had a preterm delivery (p=0.19-0.23), with the intervention group being one and a half to three times more likely to score above threshold on a depression scale (CES-D ≥16). However, the confidence in this effect estimate is very low due to risk of bias concerns (statistically significant group differences at baseline with the intervention group having more mothers with earlier preterm birth) and very serious imprecision (low event rate and 95% CI includes no effect and appreciable harm). In addition, there were contradictory effects observed for women with psychosocial risk factors, where there was single study (N=346) evidence for a

moderate effect of a mother–infant relationship intervention on preventing depression diagnosis at long-term follow-up using an available case analysis approach (p=0.22). However, this effect was not statistically or clinically significant when an ITT analysis approach was used (p=1.00), and our confidence in the effect size from the available case analysis was low due to very serious imprecision (optimal information size [events=300] was not met and 95% CI includes no effect and appreciable benefit). In addition, there was no evidence for statistically or clinically significant effects of mother–infant relationship interventions on depression diagnosis at endpoint (p=0.36-0.99), depression symptomatology at long-term follow-up (p=0.62-0.82) or on mean depression symptoms at short-term follow-up (p=0.23) or long-term follow-up (p=0.18), and no evidence for clinically significant effects on depression mean symptoms at endpoint (p=0.03).

Table 45: Summary of findings table for effects of mother–infant relationship interventions compared with treatment as usual on preventing depression outcomes in women with identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk: Control	Corresponding risk: Depression: Mother-infant relationship interventions versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Depression diagnosis post-treatment – ITT analysis (at-risk populations) SCID Follow-up: mean 26 weeks	Study population		RR 1 (0.76 to 1.31)	449 (1 study)	⊕⊕⊖⊖ low[1,2]	
	323 per 1000	323 per 1000 (246 to 423)				
	Moderate					
	323 per 1000	323 per 1000 (245 to 423)				
Depression diagnosis post-treatment – available case analysis (at-risk populations) SCID Follow-up: mean 26 weeks	Study population		RR 0.78 (0.47 to 1.32)	354 (1 study)	⊕⊕⊖⊖ low[1,2]	
	158 per 1000	123 per 1000 (74 to 208)				
	Moderate					
	158 per 1000	123 per 1000 (74 to 209)				
Depression symptomatology Post-treatment – ITT analysis (at-risk populations) CES-D ≥16 Follow-up: mean 27 weeks	Study population		RR 1.52 (0.77 to 3)	106 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	
	200 per 1000	304 per 1000 (154 to 600)				
	Moderate					
	200 per 1000	304 per 1000 (154 to 600)				

Depression symptomatology Post-treatment – available case analysis (at-risk populations) CES-D ≥16 Follow-up: mean 27 weeks	Study population		RR 2.8 (0.6 to 13.11)	87 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	
	48 per 1000	133 per 1000 (29 to 624)				
	Moderate					
	48 per 1000	134 per 1000 (29 to 629)				
Depression mean scores post-treatment – available case analysis (at-risk populations) EPDS Follow-up: 15-26 weeks		The mean depression mean scores post-treatment – available case analysis (at-risk populations) in the intervention groups was 0.22 standard deviations lower (0.41 to 0.02 lower)		417 (2 studies)	⊕⊕⊕⊕ high	SMD -0.22 (-0.41 to -0.02)
Depression mean scores Short Follow-up (9-16 weeks post-intervention) – available case analysis (at-risk populations) EPDS Follow-up: mean 28 weeks		The mean depression mean scores short follow-up (9-16 weeks post-intervention) – available case analysis (at-risk populations) in the intervention groups was 0.3 standard deviations lower (0.8 lower to 0.19 higher)		63 (1 study)	⊕⊕⊖⊖ low[2,4]	SMD -0.3 (-0.8 to 0.19)
Depression diagnosis Long Follow-up (25-103 weeks post-intervention) – ITT analysis (at-risk populations) SCID Follow-up: mean 52 weeks	Study population		RR 1 (0.77 to 1.3)	449 (1 study)	⊕⊕⊖⊖ low[1,2]	
	332 per 1000	332 per 1000 (256 to 431)				
	Moderate					
	332 per 1000	332 per 1000 (256 to 432)				
Depression diagnosis Long Follow-up (25-103 weeks post-intervention) – available case analysis (at-risk populations) SCID Follow-up: mean 52 weeks	Study population		RR 0.71 (0.41 to 1.23)	346 (1 study)	⊕⊕⊖⊖ low[1,2]	
	155 per 1000	110 per 1000 (63 to 190)				
	Moderate					
	155 per 1000	110 per 1000 (64 to 191)				
Depression symptomatology Long Follow-up (25-103 weeks post-	Study population		RR 0.94 (0.56 to 1.58)	106 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	
	360 per 1000	338 per 1000 (202 to 569)				
	Moderate					

intervention) – ITT analysis (at-risk populations) CES-D ≥16 Follow-up: mean 53 weeks	360 per 1000	338 per 1000 (202 to 569)				
Depression symptomatology Long Follow-up (25-103 weeks post-intervention) – available case analysis (at-risk populations) CES-D ≥16 Follow-up: mean 53 weeks	Study population		RR 0.75 (0.25 to 2.27)	80 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	
	158 per 1000	118 per 1000 (39 to 358)				
	Moderate					
	158 per 1000	119 per 1000 (40 to 359)				
Depression mean scores Long Follow-up (25-103 weeks post-intervention) – available case analysis (at-risk populations) EPDS Follow-up: mean 52 weeks		The mean depression mean scores long follow-up (25-103 weeks post-intervention) – available case analysis (at-risk populations) in the intervention groups was 0.14 standard deviations lower (0.35 lower to 0.06 higher)		354 (1 study)	⊕⊕⊕⊖ moderate[4]	SMD -0.14 (-0.35 to 0.06)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total number of events is less than 300 (a threshold rule-of-thumb)
[2] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
[3] Risk of bias due to statistically significant group differences at baseline
[4] Total population size is less than 400 (a threshold rule-of-thumb)

Depression: case management and individualized treatment versus treatment as usual

There was single study (N=34) evidence for a large effect (p=0.06) of case management and individualized treatment on preventing depression

symptomatology for women who had a preterm delivery or low birthweight baby (Table 46), with women in the intervention group showing a 75% risk reduction for scoring above threshold on a depression scale (BDI ≥9). However, confidence in this effect estimate is very low due to risk of bias concerns (statistically significant group differences in maternal age at baseline with older mean age in the intervention group) and very serious imprecision (with very small sample size and 95% CI including both no effect and appreciable benefit).

Table 46: Summary of findings table for effects of case management and individualized treatment compared with treatment as usual on preventing depression outcomes in women with identified risk factors

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Depression: Case management and individualized treatment versus TAU				
Depression symptomatology Post-treatment – ITT analysis (at-risk populations) BDI ≥9 Follow-up: mean 5 weeks	Study population		RR 0.25 (0.06 to 1.05)	34 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	
	438 per 1000	109 per 1000 (26 to 459)				
	Moderate					
	438 per 1000	109 per 1000 (26 to 460)				
Depression symptomatology Post-treatment -Available case analysis (at-risk populations) BDI ≥9 Follow-up: mean 5 weeks	Study population		RR 0.25 (0.06 to 1.05)	34 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	
	438 per 1000	109 per 1000 (26 to 459)				
	Moderate					
	438 per 1000	109 per 1000 (26 to 460)				

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Risk of bias due to statistically significant group differences at baseline
[2] Total number of events is less than 300 (a threshold rule-of-thumb)
[3] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

7.4.5 Clinical evidence for preventative effects on anxiety outcomes for women with identified risk factors (by intervention)

Summary of findings can be found in the tables presented in this section. The full GRADE evidence profiles and associated forest plots can be found in Appendix 22 and Appendix 19, respectively.

Anxiety: post-miscarriage self-help versus treatment as usual

There was no evidence for clinically significant effects of post-miscarriage self-help on anxiety mean symptoms, although the effect was statistically significant (p=0.0005; Table 47).

Table 47: Summary of findings table for effects of post-miscarriage self-help compared with treatment as usual on preventing anxiety outcomes in women with identified risk factors

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Anxiety: post-miscarriage self-help versus TAU				
Anxiety mean scores post-treatment – ITT analysis (at-risk populations) BSI: Anxiety Follow-up: mean 5 weeks		The mean anxiety mean scores post-treatment – ITT analysis (at-risk populations) in the intervention groups was 0.47 standard deviations lower (0.73 to 0.2 lower)		228 (1 study)	⊕⊕⊖⊖ low[1,2]	SMD -0.47 (-0.73 to -0.2)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Risk of bias due to statistically significant group differences at baseline
[2] Total population size is less than 400 (a threshold rule-of-thumb)

Anxiety: non-mental health-focused education and support versus treatment as usual or enhanced treatment as usual

There was single study (N=162) evidence for a moderate effect of non-mental health-focused education and support for preventing anxiety symptomatology (at endpoint and short-term follow-up) in women with multiple births when an ITT analysis approach was used (p=0.17-0.25) and a large effect on anxiety symptomatology at short-term follow-up when an available case analysis was used (p=0.13). However, confidence in these effect estimates was very low due to very serious imprecision (low event rate and the 95% CI includes both no effect and appreciable benefit) and selective reporting bias, and the available case analysis for anxiety symptomatology at endpoint provided no evidence for an effect on this outcome measure (p=0.89). In addition, there was no evidence for statistically or clinically significant effects on anxiety mean scores at endpoint, short-term or intermediate follow-up (p=0.14-0.34), or on anxiety symptomatology at intermediate follow-up (0.32-0.93) (Table 48).

Table 48: Summary of findings table for effects of non-mental health-focused education and support compared with treatment as usual or enhanced treatment as usual on preventing anxiety outcomes in women with identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Anxiety: Non-mental health-focused education and support versus TAU or Enhanced TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Anxiety symptomatology Post-treatment – ITT analysis (at-risk populations) HADS – Anxiety (above unspecified threshold) Follow-up: mean 6 weeks	Study population 305 per 1000	226 per 1000 (134 to 378)	RR 0.74 (0.44 to 1.24)	162 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	
	Moderate 305 per 1000	226 per 1000 (134 to 378)				
Anxiety symptomatology Post-treatment – available case analysis (at-risk	Study population 95 per 1000	89 per 1000 (30 to 259)	RR 0.93 (0.32 to 2.72)	131 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	
	Moderate					

populations) HADS – Anxiety (above unspecified threshold) Follow-up: mean 6 weeks	95 per 1000	88 per 1000 (30 to 258)				
Anxiety mean scores post-treatment – available case analysis (at-risk populations) STAI-S or HADS – Anxiety Follow-up: mean 6 weeks		The mean anxiety mean scores post-treatment – available case analysis (at-risk populations) in the intervention groups was 0.1 standard deviations lower (0.3 lower to 0.11 higher)		370 (2 studies)	⊕⊕⊕⊖ moderate[4]	SMD -0.1 (-0.3 to 0.11)
Anxiety symptomatology Short Follow-up (9-16 weeks post-intervention) – ITT analysis (at-risk populations) HADS – Anxiety (above unspecified threshold) Follow-up: mean 12 weeks	Study population		RR 0.67 (0.38 to 1.19)	162 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	
	280 per 1000	188 per 1000 (107 to 334)				
	Moderate					
	281 per 1000	188 per 1000 (107 to 334)				
Anxiety symptomatology Short Follow-up (9-16 weeks post-intervention) – available case analysis (at-risk populations) HADS – Anxiety (above unspecified threshold) Follow-up: mean 12 weeks	Study population		RR 0.11 (0.01 to 1.96)	128 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	
	63 per 1000	7 per 1000 (1 to 124)				
	Moderate					
	64 per 1000	7 per 1000 (1 to 125)				
Anxiety mean scores Short Follow-up (9-16 weeks post-intervention) – available case analysis (at-risk populations) HADS – Anxiety Follow-up: mean 12 weeks		The mean anxiety mean scores short follow-up (9-16 weeks post-intervention) – available case analysis (at-risk populations) in the intervention groups was 0.2 standard deviations lower (0.54 lower to 0.15 higher)		128 (1 study)	⊕⊖⊖⊖ very low[2,3,4]	SMD -0.2 (-0.54 to 0.15)
	Study population					

Anxiety symptomatology Intermediate follow-up (17-24 weeks post-intervention) – ITT analysis (at-risk populations) HADS – Anxiety (above unspecified threshold) Follow-up: mean 24 weeks	280 per 1000	213 per 1000 (123 to 367)	RR 0.76 (0.44 to 1.31)	162 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	
	Moderate					
	281 per 1000	214 per 1000 (124 to 368)				
Anxiety symptomatology Intermediate follow-up (17-24 weeks post-intervention) – available case analysis (at-risk populations) HADS – Anxiety (above unspecified threshold) Follow-up: mean 24 weeks	Study population		RR 0.94 (0.25 to 3.6)	130 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	
	63 per 1000	60 per 1000 (16 to 229)				
	Moderate					
	64 per 1000	60 per 1000 (16 to 230)				
Anxiety mean scores Intermediate follow-up (17-24 weeks post-intervention) – available case analysis (at-risk populations) HADS – Anxiety Follow-up: mean 24 weeks		The mean anxiety mean scores intermediate follow-up (17-24 weeks post-intervention) – available case analysis (at-risk populations) in the intervention groups was 0.26 standard deviations lower (0.6 lower to 0.09 higher)		130 (1 study)	⊕⊖⊖⊖ very low[2,3,4]	SMD -0.26 (-0.6 to 0.09)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total number of events is less than 300 (a threshold rule-of-thumb)
[2] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
[3] Paper omits data
[4] Total population size is less than 400 (a threshold rule-of-thumb)

Anxiety: home visits versus treatment as usual

There was single study (N=120) evidence for moderate to large effects of home visits on preventing anxiety symptomatology at endpoint (p=0.01) and long-term follow-up (p=0.01-0.04), and large effects observed on mean anxiety symptoms at endpoint (p <0.0001) and moderate effects on mean anxiety symptoms at long-term follow-up (p=0.009) in women who had a preterm delivery (Table 49). However, confidence in these effect estimates is very low due to risk of bias concerns (statistically significant group differences in depression symptomatology at baseline and selective reporting) and imprecision (the optimal information size [events =300/N=400] was not met).

Table 49: Summary of findings table for effects of home visits compared with treatment as usual on preventing anxiety outcomes in women with identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk / Corresponding risk Control	Anxiety: Home visits versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Anxiety symptomatology Post-treatment – ITT analysis (at-risk populations) HADS – Anxiety >7 Follow-up: mean 52 weeks	Study population		RR 0.63 (0.43 to 0.91)	120 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	
	627 per 1000	395 per 1000 (270 to 571)				
	Moderate					
	627 per 1000	395 per 1000 (270 to 571)				
Anxiety symptomatology Post-treatment – available case analysis (at-risk populations) HADS – Anxiety >7 Follow-up: mean 52 weeks	Study population		RR 0.44 (0.23 to 0.82)	90 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	
	488 per 1000	215 per 1000 (112 to 400)				
	Moderate					
	488 per 1000	215 per 1000 (112 to 400)				
Anxiety mean scores post-treatment – available case analysis (at-risk populations) HADS – Anxiety Follow-up: mean 52 weeks		The mean anxiety mean scores post-treatment – available case analysis (at-risk populations) in the intervention groups was 0.89 standard deviations lower (1.33 to 0.46 lower)		90 (1 study)	⊕⊖⊖⊖ very low[1,3,4]	SMD -0.89 (-1.33 to -0.46)
	Study population					

Anxiety symptomatology Long Follow-up (25-103 weeks post-intervention) – ITT analysis (at-risk populations) HADS – Anxiety ≥8 Follow-up: mean 104 weeks	712 per 1000	527 per 1000 (392 to 698)				
	Moderate					
	712 per 1000	527 per 1000 (392 to 698)	RR 0.74 (0.55 to 0.98)	120 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	
Anxiety symptomatology Long Follow-up (25-103 weeks post-intervention) – available case analysis (at-risk populations) HADS – Anxiety ≥8 Follow-up: mean 104 weeks	Study population		RR 0.46 (0.25 to 0.85)	77 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	
	553 per 1000	254 per 1000 (138 to 470)				
	Moderate					
	553 per 1000	254 per 1000 (138 to 470)				
Anxiety mean scores Long Follow-up (25-103 weeks post-intervention) – available case analysis (at-risk populations) HADS – Anxiety Follow-up: mean 104 weeks		The mean anxiety mean scores long follow-up (25-103 weeks post-intervention) – available case analysis (at-risk populations) in the intervention groups was 0.61 standard deviations lower (1.06 to 0.15 lower)		77 (1 study)	⊕⊖⊖⊖ very low[1,3,4]	SMD -0.61 (-1.06 to -0.15)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Risk of bias due to statistically significant group differences at baseline
[2] Total number of events is less than 300 (a threshold rule-of-thumb)
[3] Paper omits data
[4] Total population size is less than 400 (a threshold rule-of-thumb)

7.4.6 Clinical evidence for preventative effects on PTSD outcomes for women with identified risk factors (by intervention)

Summary of findings can be found in the tables presented in this section. The full GRADE evidence profiles and associated forest plots can be found in Appendix 22 and Appendix 19, respectively.

PTSD: post-miscarriage self-help versus treatment as usual

There was single study evidence (N=228) for large effects of post-miscarriage self-help on preventing PTSD symptomatology (p=0.0004) and reducing mean PTSD symptoms (p <0.00001) for women who had lost a child during pregnancy because of miscarriage, termination due to medical indications, or stillbirth (Table 50). However, confidence in these effect estimates was very low due to risk of bias concerns (statistically significant difference in baseline mean scores [lower in the intervention group] on the intrusion subscale of the IES-R) and imprecision (the optimal information size [events =300/N=400] was not met).

Table 50: Summary of findings table for effects of post-miscarriage self-help compared with treatment as usual on preventing PTSD outcomes in women with identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk PTSD: post-miscarriage self-help versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
PTSD symptomatology Post-treatment – ITT analysis (at-risk populations) IES-R ≥35 Follow-up: mean 5 weeks	Study population 310 per 1000 Moderate 310 per 1000	105 per 1000 (56 to 192) 105 per 1000 (56 to 192)	RR 0.34 (0.18 to 0.62)	228 (1 study)	⊕⊖⊖⊖ very low[1,2]	
PTSD mean scores post-treatment – ITT analysis (at-risk populations) IES-R Follow-up: mean 5 weeks		The mean PTSD mean scores post-treatment – ITT analysis (at-risk populations) in the intervention groups was 0.88 standard deviations lower (1.15 to 0.61 lower)		228 (1 study)	⊕⊖⊖⊖ very low[1,3]	SMD -0.88 (-1.15 to -0.61)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Risk of bias due to statistically significant group differences at baseline
[2] Total number of events is less than 300 (a threshold rule-of-thumb)
[3] Total population size is less than 400 (a threshold rule-of-thumb)

7.4.7 Clinical evidence for preventative effects on poor general mental health outcomes for women with identified risk factors (by intervention)

Summary of findings can be found in the tables presented in this section. The full GRADE evidence profiles and associated forest plots can be found in Appendix 22 and Appendix 19, respectively.

General mental health: post-miscarriage self-help versus treatment as usual

There was single study evidence (N=228) for a moderate benefit of post-miscarriage self-help on preventing poor general mental health outcomes ($p < 0.00001$) for women who had lost a child during pregnancy because of miscarriage, termination due to medical indications, or stillbirth. However, the confidence in this effect estimate was low due to risk of bias concerns (statistically significant group difference at baseline) and small sample size (Table 51).

Table 51: Summary of findings table for effects of post-miscarriage self-help compared with treatment as usual on preventing poor general mental health outcomes in women with identified risk factors

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	General mental health: post-miscarriage self-help versus TAU				
General mental health mean scores post-treatment – ITT		The mean general mental health mean scores post-treatment – ITT analysis (at-risk		228 (1 study)	⊕⊕⊖⊖ low[1,2]	SMD -0.61 (-0.87 to -0.34)

analysis (at-risk populations) BSI: Global severity index (Mental health) Follow-up: mean 5 weeks	populations) in the intervention groups was 0.61 standard deviations lower (0.87 to 0.34 lower)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] High risk of selection bias due to unclear allocation concealment and statistically significant difference in baseline intrusion subscale of the IES-R (19.2 in control group and 17.4 in intervention group)
[2] Total population size is less than 400 (a threshold rule-of-thumb)

General mental health: home visits versus treatment as usual

Two studies (N=207) provided no evidence for a clinically or statistically significant effect of home visits on preventing poor general mental health outcomes (p=0.49) in women with psychosocial risk factors and who were adolescent or had a (family) history of mental health problems (Table 52).

Table 52: Summary of findings table for effects of home visits compared with treatment as usual on preventing poor general mental health outcomes in women with identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	 Corresponding risk General mental health: Home visits versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
General mental health mean scores post-treatment – available case analysis (at-risk populations) General Health Questionnaire		The mean general mental health mean scores post-treatment – available case analysis (at-risk populations) in the intervention groups was 0.18 standard deviations		207 (2 studies)	⊕⊖⊖⊖ very low[1,2,3,4]	SMD -0.18 (-0.7 to 0.33)

(GHQ) Follow-up: mean 78 weeks	lower (0.7 lower to 0.33 higher)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] There is evidence of substantial heterogeneity of study effect sizes
[2] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
[3] Total population size is less than 400 (a threshold rule-of-thumb)
[4] Paper omits data

General mental health: post-delivery discussion versus enhanced treatment as usual

A single study (N=534-917) failed to find evidence for clinically or statistically significant benefits of a midwife-led post-delivery discussion relative to a non-mental health-focused information booklet on preventing poor general mental health outcomes at post-treatment (p=0.22) or very long (208-312 weeks) follow-up (p=0.05) for women who had had an operative delivery (Table 53).

Table 53: Summary of findings table for effects of post-delivery discussion compared with enhanced treatment as usual on preventing poor general mental health outcomes in women with identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	 Corresponding risk General mental health: post-delivery discussion versus Enhanced TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
General mental health mean scores post-treatment – available case analysis (at-risk populations) SF-36- Mental health		The mean general mental health mean scores post-treatment – available case analysis (at-risk populations) in the intervention groups was 0.08 standard deviations		917 (1 study)	⊕⊕⊕⊕ high	SMD -0.08 (-0.21 to 0.05)

Follow-up: mean 26 weeks	lower (0.21 lower to 0.05 higher)			
General mental health mean scores Very long follow-up (>104 weeks post-intervention) – available case analysis (at-risk populations) SF-36- Mental health Follow-up: 208-312 weeks	The mean general mental health mean scores very long follow-up (>104 weeks post-intervention) – available case analysis (at-risk populations) in the intervention groups was 0.17 standard deviations higher (0 to 0.34 higher)	534 (1 study)	⊕⊕⊕⊕ high	SMD 0.17 (0 to 0.34)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

General mental health: mother-infant relationship interventions versus treatment as usual

A single study (N=88-125) found no evidence for clinically or statistically significant benefits of a mother–infant relationship intervention relative to treatment as usual on preventing poor general mental health outcomes at post-treatment (p=0.31) or long follow-up (p=0.66) for women who had a preterm delivery or a baby with low birthweight (Table 54).

Table 54: Summary of findings table for effects of mother–infant relationship interventions compared with treatment as usual on preventing poor general mental health outcomes in women with identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk General mental health: Mother-infant relationship interventions versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
General mental health mean scores post-treatment – available case analysis (at-risk populations) GHQ-28 Follow-up: mean 26 weeks		The mean general mental health mean scores post-treatment – available case analysis (at-risk populations) in the intervention groups was 0.18 standard deviations higher (0.17 lower to 0.53 higher)		125 (1 study)	⊕⊕⊖⊖ low[1,2]	SMD 0.18 (-0.17 to 0.53)
General mental health mean scores Long follow-up (25-104 weeks post-intervention) – available case analysis (at-risk populations) GHQ-28 Follow-up: mean 104 weeks		The mean general mental health mean scores long follow-up (25-104 weeks post-intervention) – available case analysis (at-risk populations) in the intervention groups was 0.09 standard deviations lower (0.52 lower to 0.33 higher)		88 (1 study)	⊕⊕⊖⊖ low[1,2]	SMD -0.09 (-0.52 to 0.33)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total population size is less than 400 (a threshold rule-of-thumb)
[2] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

7.4.8 Clinical evidence for preventative effects on poor mental health outcomes for women with identified risk factors (sub-analyses)

There was insufficient data to enable sub-analyses by risk factor, treatment timing, format or intensity for the prevention (risk factors identified) review.

7.4.9 Clinical evidence for preventative effects on mother–infant attachment problems for women with identified risk factors (by intervention)

Summary of findings can be found in the tables presented in this section. The full GRADE evidence profiles and associated forest plots can be found in Appendix 22 and Appendix 19, respectively.

Mother-infant attachment: non-mental health-focused education and support versus treatment as usual or enhanced treatment as usual

A single study (N=126) found evidence for a moderate harm of non-mental health-focused education and support group and home visits relative to treatment as usual at short follow-up (p=0.32) for women with an uncomplicated twin pregnancy when an available case analysis approach was used (Table 55). However, confidence in this effect estimate was very low due to very serious imprecision (number of events fell below the threshold rule-of-thumb for optimal information size and the 95% CI included both no effect and measures of appreciable harm) and risk of selective reporting bias. This study (N=162) found no evidence for a clinically or statistically significant effect on this outcome measure at this time point when an ITT analysis approach was used (p=0.64). Moreover, no clinically or statistically significant effects were observed at post-treatment (N=133-162; p=0.52-0.97) or at intermediate follow-up (N=127-162; p=0.28-0.58).

Another single study (N=199-241) found evidence for small to moderate benefits of a non-mental health-focused education and support (booklet and audiotaped) intervention on preventing poor mother–infant interaction mean scores (p<0.0001) or poor maternal sensitivity (p=0.04) for mothers with babies in the neonatal intensive care unit (NICU) who had had preterm delivery and low birthweight babies (Table 55). However, confidence in these effect estimates was low to very low due to imprecision and selective reporting bias. This study found no evidence for a clinically or statistically significant effect of non-mental health-focused education and support on preventing poor maternal confidence (p=0.24).

Table 55: Summary of findings table for effects of non-mental health-focused education and support compared with treatment as usual or enhanced treatment as usual on preventing mother–infant attachment problems for women with identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk / Corresponding risk Control	Mother-infant attachment: Non-mental health-focused education and support versus TAU or Enhanced TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Mother-infant attachment problems Post-treatment – ITT analysis (at-risk populations) Green scale: Mother-infant attachment problems (above unspecified threshold) Follow-up: mean 6 weeks	Study population		RR 0.9 (0.65 to 1.25)	162 (1 study)	⊕⊖⊖⊖ very low1,2,3	
	500 per 1000	450 per 1000 (325 to 625)				
	Moderate					
	500 per 1000	450 per 1000 (325 to 625)				
Mother-infant attachment problems Post-treatment – available case analysis (at-risk populations) Green scale: Mother-infant attachment problems (above unspecified threshold) Follow-up: mean 6 weeks	Study population		RR 1.01 (0.64 to 1.59)	133 (1 study)	⊕⊖⊖⊖ very low1,2,3	
	359 per 1000	363 per 1000 (230 to 571)				
	Moderate					
	359 per 1000	363 per 1000 (230 to 571)				
Positive mother–infant interaction mean scores post-treatment – available case analysis (at-risk populations) Index of Parental Behavior in the NICU: Positive		The mean positive mother–infant interaction mean scores post-treatment – available case analysis (at-risk populations) in the intervention groups was 0.57 standard		211 (1 study)	⊕⊕⊖⊖ low3,4	SMD 0.57 (0.29 to 0.85)

interaction with quiet alert infant		deviations higher (0.29 to 0.85 higher)				
Maternal sensitivity mean scores post-treatment – available case analysis (at-risk populations) Index of Parental Behavior in the NICU: Sensitivity to needs of infant in NICU		The mean maternal sensitivity mean scores post-treatment – available case analysis (at-risk populations) in the intervention groups was 0.3 standard deviations higher (0.02 to 0.58 higher)		199 (1 study)	⊕⊖⊖⊖ very low3,4	SMD 0.3 (0.02 to 0.58)
Maternal confidence mean scores post-treatment – available case analysis (at-risk populations) Parental Belief Scale-NICU: Parent role confidence		The mean maternal confidence mean scores post-treatment – available case analysis (at-risk populations) in the intervention groups was 0.15 standard deviations higher (0.1 lower to 0.41 higher)		241 (1 study)	⊕⊕⊖⊖ low3,4	SMD 0.15 (-0.1 to 0.41)
Mother-infant attachment problems Short Follow-up (9-16 weeks post-intervention) – ITT analysis (at-risk populations) Green scale: Mother-infant attachment problems (above unspecified threshold) Follow-up: mean 12 weeks	Study population 463 per 1000 Moderate 463 per 1000	500 per 1000 (361 to 690) 500 per 1000 (361 to 690)	RR 1.08 (0.78 to 1.49)	162 (1 study)	⊕⊖⊖⊖ very low1,2,3	
Mother-infant attachment problems Short Follow-up (9-16 weeks post-intervention) – available case analysis (at-risk populations) Green scale: Mother-infant attachment problems (above unspecified threshold) Follow-up: mean 12 weeks	Study population 290 per 1000 Moderate 290 per 1000	375 per 1000 (226 to 618) 374 per 1000 (226 to 618)	RR 1.29 (0.78 to 2.13)	126 (1 study)	⊕⊖⊖⊖ very low1,2,3	

Mother-infant attachment problems Intermediate follow-up (17-24 weeks post-intervention) – ITT analysis (at-risk populations) Green scale: Mother-infant attachment problems (above unspecified threshold) Follow-up: mean 24 weeks	Study population 585 per 1000	498 per 1000 (375 to 667)	RR 0.85 (0.64 to 1.14)	162 (1 study)	⊕⊖⊖⊖ very low1,2,3
	Moderate 585 per 1000	497 per 1000 (374 to 667)			
Mother-infant attachment problems Intermediate follow-up (17-24 weeks post-intervention) – available case analysis (at-risk populations) Green scale: Mother-infant attachment problems (above unspecified threshold) Follow-up: mean 24 weeks	Study population 443 per 1000	394 per 1000 (261 to 593)	RR 0.89 (0.59 to 1.34)	127 (1 study)	⊕⊖⊖⊖ very low1,2,3
	Moderate 443 per 1000	394 per 1000 (261 to 594)			

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total number of events is less than 300 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
3 Paper omits data
4 Total population size is less than 400 (a threshold rule-of-thumb)

Mother-infant attachment: home visits versus treatment as usual

There was single study (N=121-131) evidence for small and statistically significant benefits of home visits relative to treatment as usual for preventing poor maternal sensitivity (p=0.05) or poor infant involvement (p=0.02) for women with

psychosocial risk factors and (family) history of mental health problems. However, these estimates did not meet the criteria for clinically appreciable benefits and confidence in the effect estimates was very low due to very serious imprecision and selective reporting bias (Table 56). This same study found no evidence for clinically or statistically significant effects of home visits on preventing the discontinuation of breastfeeding before 6 months (p=0.30).

Table 56: Summary of findings table for effects of home visits compared with treatment as usual on preventing mother–infant attachment problems for women with identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Mother-infant attachment: Home visits versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Maternal sensitivity mean scores post-treatment – available case analysis (at-risk populations) CARE Index scale – Maternal sensitivity Follow-up: mean 78 weeks		The mean maternal sensitivity mean scores post-treatment – available case analysis (at-risk populations) in the intervention groups was 0.36 standard deviations higher (0 to 0.72 higher)		121 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	SMD 0.36 (0 to 0.72)
Infant involvement mean scores post-treatment – available case analysis (at-risk populations) CARE Index scale – Infant cooperativeness Follow-up: mean 78 weeks		The mean infant involvement mean scores post-treatment – available case analysis (at-risk populations) in the intervention groups was 0.42 standard deviations higher (0.06 to 0.78 higher)		121 (1 study)	⊕⊖⊖⊖ very low[1,3]	SMD 0.42 (0.06 to 0.78)
Discontinued breastfeeding <6 months – ITT analysis (at-risk populations) Breastfeeding – discontinued before 6 months Follow-up: mean 52 weeks	Study population 381 per 1000 Moderate 381 per 1000	293 per 1000 (183 to 476) 293 per 1000 (183 to 476)	RR 0.77 (0.48 to 1.25)	131 (1 study)	⊕⊖⊖⊖ very low[2,3,4]	

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the

estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total population size is less than 400 (a threshold rule-of-thumb)
[2] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
[3] Paper omits data
[4] Total number of events is less than 300 (a threshold rule-of-thumb)

Mother-infant attachment: mother-infant relationship interventions versus treatment as usual

There was single study (N=318-449) low quality evidence for a moderate benefit of a mother–infant relationship intervention on preventing mother–infant attachment problems in women with psychosocial risk factors when an available case analysis approach was used (p=0.03). However, this effect was not clinically or statistically significant when an ITT (WCS) analysis approach was adopted (p=0.08). There was also evidence from two studies (N=172-175) for a small benefit of mother–infant relationship interventions on preventing poor mother–infant interaction mean scores (p=0.003) for women who had had a preterm delivery and/or a low birthweight baby. However, this effect estimate did not reach criteria for a clinically meaningful benefit (SMD<0.5), only available case analysis was reported, and confidence in the effect estimate was low as the sample size was below the threshold rule-of-thumb for the optimal information size (N=400). There was also evidence from the same two studies for moderate effects of mother–infant relationship interventions on preventing poor maternal sensitivity (p=0.10) and infant responsivity (p=0.38) mean scores. However, these effects were not statistically significant and the evidence was very low quality due to very serious imprecision and considerable heterogeneity (I^2=80-92%). Single study analyses (N=109-112) failed to find evidence for clinically or statistically significant effects of mother–infant relationship interventions on preventing poor maternal intrusiveness (p=0.10), infant involvement (p=0.10) or infant negative engagement/behaviour problems (p=0.40) mean scores and effect size could not be estimated for maternal negative engagement due to zero count cells (Table 57).

Another single study (N=81-106) found evidence for clinically significant, or clinically and statistically significant, benefits of a mother–infant relationship intervention for preventing breastfeeding discontinuation before 6 months (p=0.17) or 9 months (p=0.03) for women who had had a preterm delivery when an available case analysis approach was used (Table 57). However, the quality of the evidence was very low and there was no evidence for clinically or statistically significant effects when an ITT analysis approach was used for preventing breastfeeding discontinuation before 6 months (p=0.62) or 9 months (p=0.09), and no clinically or statistically significant effects were observed for preventing breastfeeding discontinuation before 12 months when either an available case (p=0.08) or an ITT (p=0.12) analysis approach was used.

Table 57: Summary of findings table for effects of mother–infant relationship interventions compared with treatment as usual on preventing mother–infant attachment problems for women with identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Mother-infant attachment: Mother-infant relationship interventions versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Mother-infant attachment problems Post-treatment – ITT analysis (at-risk populations) Ainsworth Strange Situation: Insecure Follow-up: mean 78 weeks	Study population		RR 0.85 (0.71 to 1.02)	449 (1 study)	⊕⊕⊖⊖ low[1,2]	
	555 per 1000	471 per 1000 (394 to 566)				
	Moderate					
	555 per 1000	472 per 1000 (394 to 566)				
Mother-infant attachment problems Post-treatment – available case analysis (at-risk populations) Ainsworth Strange Situation: Insecure Follow-up: mean 78 weeks	Study population		RR 0.69 (0.5 to 0.97)	318 (1 study)	⊕⊕⊖⊖ low[1]	
	370 per 1000	256 per 1000 (185 to 359)				
	Moderate					
	370 per 1000	255 per 1000 (185 to 359)				
Positive mother–infant interaction mean scores post-treatment – available case analysis (at-risk populations) Infant and Caregiver Engagement Phases (ICEP): Maternal positive engagement (% of time during behavioural observation) or Synchrony Scale (Milgrom & Meitz, 1988): Reciprocity/Synchrony Follow-up: 15-26 weeks		The mean positive mother–infant interaction mean scores post-treatment – available case analysis (at-risk populations) in the intervention groups was 0.46 standard deviations higher (0.16 to 0.76 higher)		175 (2 studies)	⊕⊕⊖⊖ low[3]	SMD 0.46 (0.16 to 0.76)
Maternal sensitivity mean scores post-treatment – available case analysis (at-risk populations) Maternal Sensitivity and Responsivity Scales: Maternal sensitivity or synchrony (Milgrom & Meitz, 1988): Maternal Respond Follow-up: 15-26 weeks		The mean maternal sensitivity mean scores post-treatment – available case analysis (at-risk populations) in the intervention groups was 0.62 standard deviations higher		172 (2 studies)	⊕⊖⊖⊖ very low[2,3,4]	SMD 0.62 (-0.11 to 1.35)

		(0.11 lower to 1.35 higher)				
Maternal intrusiveness mean scores post-treatment – available case analysis (at-risk populations) Maternal Sensitivity and Responsivity Scales: Maternal intrusiveness Follow-up: mean 26 weeks		The mean maternal intrusiveness mean scores post-treatment – available case analysis (at-risk populations) in the intervention groups was 0.32 standard deviations lower (0.7 lower to 0.06 higher)		109 (1 study)	⊕⊕⊖⊖ low[2,3]	SMD -0.32 (-0.7 to 0.06)
Maternal negative engagement mean scores post-treatment – available case analysis (at-risk populations) ICEP: Maternal negative engagement (angry/hostile/stern/sad/sober/ expressionless; % of time during behavioural observation) Follow-up: mean 26 weeks	See comment	See comment	Not estimable	112 (1 study)	⊕⊕⊖⊖ low[3]	
Infant involvement mean scores post-treatment – available case analysis (at-risk populations) ICEP: Infant positive engagement (% of time during behavioural observation) Follow-up: mean 26 weeks		The mean infant involvement mean scores post-treatment – available case analysis (at-risk populations) in the intervention groups was 0.31 standard deviations lower (0.69 lower to 0.06 higher)		112 (1 study)	⊕⊕⊖⊖ low[2,3]	SMD -0.31 (-0.69 to 0.06)
Infant responsivity mean scores post-treatment – available case analysis (at-risk populations) ICEP: Infant responsivity (mother-focused attention; % of time during behavioural observation) or Synchrony Scale (Milgrom & Meitz, 1988): Attending to mother Follow-up: 15-26 weeks		The mean infant responsivity mean scores post-treatment – available case analysis (at-risk populations) in the intervention groups was 0.52 standard deviations higher (0.63 lower to 1.68 higher)		175 (2 studies)	⊕⊖⊖⊖ very low[2,3,4]	SMD 0.52 (-0.63 to 1.68)
Infant negative engagement/behaviour problems mean score Post-treatment – available case analysis (at-risk populations)		The mean infant negative engagement/behaviour problems mean score post-treatment		112 (1 study)	⊕⊕⊖⊖ low[2,3]	SMD 0.16 (-0.21 to 0.53)

ICEP: Infant negative engagement (behaviour problems; % of time during behavioural observation) Follow-up: mean 26 weeks	– available case analysis (at-risk populations) in the intervention groups was 0.16 standard deviations higher (0.21 lower to 0.53 higher)			
Discontinued breastfeeding <6 months – ITT analysis (at-risk populations) Infant feeding-breast feeding stopped by 26 weeks Follow-up: mean 27 weeks	Study population 440 per 1000 — 392 per 1000 (251 to 616) Moderate 440 per 1000 — 392 per 1000 (251 to 616)	RR 0.89 (0.57 to 1.4)	106 (1 study)	⊕⊖⊖⊖ very low[1,2,5,6]
Discontinued breastfeeding <6 months – available case analysis (at-risk populations) Infant feeding-breast feeding stopped by 26 weeks Follow-up: mean 27 weeks	Study population 364 per 1000 — 225 per 1000 (116 to 444) Moderate 364 per 1000 — 226 per 1000 (116 to 444)	RR 0.62 (0.32 to 1.22)	88 (1 study)	⊕⊖⊖⊖ very low[1,2,5,6]
Discontinued breastfeeding <9 months – ITT analysis (at-risk populations) Infant feeding-breast feeding stopped by 39 weeks Follow-up: mean 40 weeks	Study population 680 per 1000 — 517 per 1000 (381 to 707) Moderate 680 per 1000 — 517 per 1000 (381 to 707)	RR 0.76 (0.56 to 1.04)	106 (1 study)	⊕⊖⊖⊖ very low[1,2,5,6]
Discontinued breastfeeding <9 months – available case analysis (at-risk populations) Infant feeding-breast feeding stopped by 39 weeks Follow-up: mean 40 weeks	Study population 600 per 1000 — 342 per 1000 (210 to 558) Moderate 600 per 1000 — 342 per 1000 (210 to 558)	RR 0.57 (0.35 to 0.93)	81 (1 study)	⊕⊖⊖⊖ very low[1,5,6]
Discontinued breastfeeding <12 months – ITT analysis (at-risk populations) Infant feeding-breast feeding stopped by 52 weeks Follow-up: mean 53 weeks	Study population 840 per 1000 — 714 per 1000 (580 to 874) Moderate 840 per 1000 — 714 per 1000 (580 to 874)	RR 0.85 (0.69 to 1.04)	106 (1 study)	⊕⊖⊖⊖ very low[1,2,5,6]
Discontinued breastfeeding <12 months – available case analysis (at-risk populations) Infant feeding-breast feeding stopped by 52 weeks Follow-up: mean 53 weeks	Study population 800 per 1000 — 616 per 1000 (464 to 824) Moderate 800 per 1000 — 616 per 1000 (464 to 824)	RR 0.77 (0.58 to 1.03)	82 (1 study)	⊕⊖⊖⊖ very low[1,2,5,6]

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the

comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total number of events is less than 300 (a threshold rule-of-thumb)
[2] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
[3] Total population size is less than 400 (a threshold rule-of-thumb)
[4] There is evidence of considerable heterogeneity of study effect sizes
[5] High risk of selection bias due to statistically significant baseline difference with the intervention group having more mothers with earlier preterm birth and non-Norwegian origin
[6] Paper omits data

Mother-infant attachment: case management and individualized treatment versus treatment as usual

There was single study (N=30) very low quality evidence for a moderate benefit of case management and individualized treatment on preventing maternal sensitivity problems (p=0.08) for women who had had a preterm delivery and low birthweight baby (Table 58). However, this effect was not statistically significant due to very serious imprecision and there was a high risk of selection bias due to statistically significant group differences at baseline.

Table 58: Summary of findings table for effects of case management and individualized treatment compared with treatment as usual on preventing mother–infant attachment problems for women with identified risk factors

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Mother-infant attachment: Case management and individualized treatment versus TAU				
Maternal sensitivity Post-treatment – ITT analysis (at-risk populations)	Study population		RR 1.4 (0.95 to 2.05)	30 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	
	667 per 1000	933 per 1000 (633 to 1000)				
	Moderate					

Behavioural observation: Maternal sensitivity Follow-up: mean 5 weeks	667 per 1000	934 per 1000 (634 to 1000)			
Maternal sensitivity Post-treatment – available case analysis (at-risk populations) Behavioural observation: Maternal sensitivity Follow-up: mean 5 weeks	Study population 667 per 1000 Moderate 667 per 1000	933 per 1000 (633 to 1000) 934 per 1000 (634 to 1000)	RR 1.4 (0.95 to 2.05)	30 (1 study)	⊕⊖⊖⊖ very low[1,2,3]

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] High risk of selection bias due to unclear allocation concealment and statistically significant baseline difference in maternal age (29.7 in intervention group and 25.9 in control group)
[2] Total number of events is less than 300 (a threshold rule-of-thumb)
[3] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

7.4.10 Clinical evidence for preventative effects on poor quality of life outcomes for women with identified risk factors (by intervention)

Summary of findings can be found in the tables presented in this section. The full GRADE evidence profiles and associated forest plots can be found in Appendix 22 and Appendix 19, respectively.

Quality of life: Psychologically (CBT/IPT)-informed psychoeducation versus treatment as usual or enhanced treatment as usual

A single study (N=190-209) found no evidence for clinically or statistically significant effects of CBT-informed psychoeducation relative to treatment as usual on preventing poor social support (p=0.61-0.78) for pregnant women with psychosocial risk factors (Table 59).

Table 59: Summary of findings table for effects of psychologically (CBT/IPT)-informed psychoeducation compared with treatment as usual or enhanced treatment as usual on preventing poor quality of life outcomes for women with identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Quality of life: Psychologically (CBT/IPT)-informed psychoeducation versus TAU or Enhanced TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Poor social support Post-treatment – ITT analysis (at-risk populations) Poor social support (interview) Follow-up: mean 27 weeks	Study population		RR 1.08 (0.62 to 1.87)	209 (1 study)	⊕⊕⊖⊖ low[1,2]	
	189 per 1000	204 per 1000 (117 to 353)				
	Moderate					
	189 per 1000	204 per 1000 (117 to 353)				
Poor social support Post-treatment – available case (at-risk populations) Poor social support (interview) Follow-up: mean 27 weeks	Study population		RR 1.23 (0.56 to 2.7)	190 (1 study)	⊕⊕⊖⊖ low[1,2]	
	104 per 1000	128 per 1000 (58 to 281)				
	Moderate					
	104 per 1000	128 per 1000 (58 to 281)				

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total number of events is less than 300 (a threshold rule-of-thumb)
[2] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Quality of life: non-mental health-focused education and support versus treatment as usual or enhanced treatment as usual

There was low quality evidence from two studies (N=369) for a small benefit of non-mental health-focused education and support (booklet and audiotaped or support group and home visits) on preventing high maternal stress (p=0.002) in women who had had a preterm delivery and low birthweight baby or women who had an uncomplicated twin pregnancy (Table 60). However, the threshold rule-of-thumb for the optimal information size (N=400) was not met and there was a high risk of selective reporting bias. Single study analyses (N=127-133) found very low quality evidence for a clinically and statistically significant benefit of a non-mental health-focused education and support group and home visits relative to treatment as usual on preventing poor social support at intermediate follow-up (p=0.004), a statistically but not clinically significant benefit at short-term follow-up (p=0.03), and no evidence of clinically or statistically significant benefits at post-treatment (p=0.20) for women with an uncomplicated twin pregnancy.

Table 60: Summary of findings table for effects of non-mental health-focused education and support compared with treatment as usual or enhanced treatment as usual on preventing poor quality of life outcomes for women with identified risk factors

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Quality of life: Non-mental health-focused education and support versus TAU or Enhanced TAU				
Parental stress mean scores post-treatment – available case analysis (at-risk populations) Parental Stressor Scale-Neonatal Intensive Care (PSS-NICU) or Parenting Stress Index (PSI)		The mean parental stress mean scores post-treatment – available case analysis (at-risk populations) in the intervention groups was 0.44 standard deviations lower (0.72 to 0.16 lower)		369 (2 studies)	⊕⊕⊖⊖ low1,2	SMD -0.44 (-0.72 to -0.16)

Follow-up: 0.4-24 weeks				
Social support mean scores post-treatment – available case analysis (at-risk populations) Satisfaction with Motherhood scale: Social support Follow-up: mean 6 weeks	The mean social support mean scores post-treatment – available case analysis (at-risk populations) in the intervention groups was 0.22 standard deviations higher (0.12 lower to 0.57 higher)	133 (1 study)	⊕⊖⊖⊖ very low1,2,3	SMD 0.22 (-0.12 to 0.57)
Social support mean scores Short Follow-up (9-16 weeks post-intervention) – available case analysis (at-risk populations) Satisfaction with Motherhood scale: Social support Follow-up: mean 12 weeks	The mean social support mean scores short follow-up (9-16 weeks post-intervention) – available case analysis (at-risk populations) in the intervention groups was 0.39 standard deviations higher (0.04 to 0.74 higher)	127 (1 study)	⊕⊖⊖⊖ very low1,2	SMD 0.39 (0.04 to 0.74)
Social support mean scores Intermediate follow-up (17-24 weeks post-intervention) – available case analysis (at-risk populations) Satisfaction with Motherhood scale: Social support Follow-up: mean 24 weeks	The mean social support mean scores intermediate follow-up (17-24 weeks post-intervention) – available case analysis (at-risk populations) in the intervention groups was 0.52 standard deviations higher (0.17 to 0.87 higher)	129 (1 study)	⊕⊖⊖⊖ very low1,2	SMD 0.52 (0.17 to 0.87)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total population size is less than 400 (a threshold rule-of-thumb)
2 Papers omit data

3 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Quality of life: home visits versus treatment as usual

There was single study (N=29) evidence for a moderate benefit of home visits relative to treatment as usual for preventing poor social support (p=0.13) for women with psychosocial risk factors and (family) history of mental health problems (Table 61). However, this effect was not statistically significant due to very serious imprecision and there was a high risk of selective reporting bias. The same study (N=114) found no evidence for clinically or statistically significant benefits of home visits on preventing poor self-esteem (p=0.83).

Table 61: Summary of findings table for effects of home visits compared with treatment as usual on preventing poor quality of life outcomes for women with identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Quality of life: Home visits versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Social support mean scores post-treatment – available case analysis (at-risk populations) Social Support Questionnaire Follow-up: mean 78 weeks		The mean social support mean scores post-treatment – available case analysis (at-risk populations) in the intervention groups was 0.58 standard deviations higher (0.17 lower to 1.34 higher)		29 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	SMD 0.58 (-0.17 to 1.34)
Self-esteem mean scores post-treatment – available case analysis (at-risk populations) Rosenberg Self-Esteem Scale (SES) Follow-up: mean 78 weeks		The mean self-esteem mean scores post-treatment – available case analysis (at-risk populations) in the intervention groups was 0.04 standard deviations lower (0.41 lower to 0.33 higher)		114 (1 study)	⊕⊖⊖⊖ very low[1,3]	SMD -0.04 (-0.41 to 0.33)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total population size is less than 400 (a threshold rule-of-thumb)
[2] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
[3] Paper omits data

Quality of life: mother-infant relationship interventions versus treatment as usual

Two to three studies (N=183-244) found no evidence for clinically or statistically significant effects of mother–infant relationship interventions on preventing high parental stress at post-treatment (p=0.21) or long follow-up (p=0.92) for women who had had a preterm delivery and/or low birthweight baby (Table 62).

Table 62: Summary of findings table for effects of mother–infant relationship interventions compared with treatment as usual on preventing poor quality of life outcomes for women with identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Quality of life: Mother-infant relationship interventions versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Parental stress mean scores post-treatment – available case analysis (at-risk populations) Nijmeegse Ouderlijke Stress Index or PSI Follow-up: 15-52 weeks		The mean parental stress mean scores post-treatment – available case analysis (at-risk populations) in the intervention groups was **0.16 standard deviations higher** (0.09 lower to 0.41 higher)		244 (3 studies)	⊕⊕⊕⊖ **moderate**[1]	SMD 0.16 (-0.09 to 0.41)
Parental stress mean scores Long follow-up (25-104 weeks post-intervention) – available case		The mean parental stress mean scores long follow-up (25-104 weeks post-intervention) – available case analysis		183 (2 studies)	⊕⊕⊖⊖ **low**[1]	SMD -0.02 (-0.33 to 0.29)

analysis (at-risk populations) Nijmeegse Ouderlijke Stress Index or PSI Follow-up: 53-104 weeks	(at-risk populations) in the intervention groups was **0.02 standard deviations lower** (0.33 lower to 0.29 higher)

*The basis for the **assumed risk** (for example, the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% CI) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total population size is less than 400 (a threshold rule-of-thumb)

Quality of life: case management and individualized treatment versus treatment as usual

A single study (N=34) found no evidence for clinically or statistically significant benefits of case management and individualized treatment relative to treatment as usual for preventing high maternal stress (p=0.22) or poor self-esteem (p=0.39) for women who have had a preterm delivery and low birthweight baby (Table 63).

Table 63: Summary of findings table for effects of case management and individualized treatment compared with treatment as usual on preventing poor quality of life outcomes for women with identified risk factors

Outcomes	**Illustrative comparative risks* (95% CI)** **Assumed risk** **Control**	**Corresponding risk** **Quality of life: Case management and individualized treatment versus TAU**	**Relative effect (95% CI)**	**No. of Participants (studies)**	**Quality of the evidence (GRADE)**	**Comments**
Parental stress mean scores post-treatment – ITT analysis (at-risk populations) PSS-NICU Follow-up: mean 5 weeks		The mean parental stress mean scores post-treatment – ITT analysis (at-risk populations) in the intervention groups was 0.43 standard deviations lower		34 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	SMD -0.43 (-1.11 to 0.25)

	(1.11 lower to 0.25 higher)			
Parental stress mean scores post-treatment – available case analysis (at-risk populations) – Case management and individualized treatment PSS-NICU Follow-up: mean 5 weeks	The mean parental stress mean scores post-treatment – available case analysis (at-risk populations) – case management and individualized treatment in the intervention groups was 0.43 standard deviations lower (1.11 lower to 0.25 higher)	34 (1 study)	⊕⊖⊖⊖ very low1,2,3	SMD -0.43 (-1.11 to 0.25)
Self-esteem mean scores post-treatment – ITT analysis (at-risk populations) Maternal Self-Report Inventory Follow-up: mean 5 weeks	The mean self-esteem mean scores post-treatment – ITT analysis (at-risk populations) in the intervention groups was 0.3 standard deviations lower (0.97 lower to 0.38 higher)	34 (1 study)	⊕⊖⊖⊖ very low1,2,3	SMD -0.3 (-0.97 to 0.38)
Self-esteem mean scores post-treatment – available case analysis (at-risk populations) Maternal Self-Report Inventory Follow-up: mean 5 weeks	The mean self-esteem mean scores post-treatment – available case analysis (at-risk populations) in the intervention groups was 0.3 standard deviations lower (0.97 lower to 0.38 higher)	34 (1 study)	⊕⊖⊖⊖ very low1,2,3	SMD -0.3 (-0.97 to 0.38)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 High risk of selection bias due to unclear allocation concealment and statistcally significant baseline difference in maternal age (29.7 in intervention group and 25.9 in control group)
2 Total population size is less than 400 (a threshold rule-of-thumb)
3 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

7.4.11 Clinical evidence for preventative effects on service utilisation for women with identified risk factors (by intervention)

Summary of findings can be found in the tables presented in this section. The full GRADE evidence profiles and associated forest plots can be found in Appendix 22 and Appendix 19, respectively.

Service utilisation: Psychologically (CBT/IPT)-informed psychoeducation versus treatment as usual or enhanced treatment as usual

A single study (N=190-209) found no evidence for clinically or statistically significant effects of CBT-informed psychoeducation relative to treatment as usual for preventing poor service utilisation (p=0.61-0.62) for women with psychosocial risk factors (Table 64).

Table 64: Summary of findings table for effects of psychologically (CBT/IPT)-informed psychoeducation compared with treatment as usual or enhanced treatment as usual on preventing poor service utilisation for women with identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Service utilisation: Psychologically (CBT/IPT)-informed psychoeducation versus TAU or Enhanced TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Contact with primary and/or secondary care Post-Treatment – ITT analysis (at-risk populations) Primary and secondary health service contact since randomisation Follow-up: mean 27 weeks	Study population 104 per 1000	127 per 1000 (59 to 269)	RR 1.22 (0.57 to 2.59)	209 (1 study)	⊕⊕⊖⊖ low[1,2]	
	Moderate 104 per 1000	127 per 1000 (59 to 269)				
Contact with primary and/or secondary care Post-treatment – available case analysis (at-risk populations) Primary and	Study population 115 per 1000	139 per 1000 (65 to 293)	RR 1.21 (0.57 to 2.56)	190 (1 study)	⊕⊕⊖⊖ low[1,2]	
	Moderate 115 per 1000	139 per 1000 (66 to 294)				

secondary health service contact since randomisation Follow-up: mean 27 weeks	

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total number of events is less than 300 (a threshold rule-of-thumb)
[2] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Service utilisation: home visits versus treatment as usual

A single study (N=63) found very low quality evidence for a moderate benefit of home visits on preventing poor maternal contact with primary and/or secondary care for adolescent women with psychosocial risk factors when an available case analysis was adopted (p=0.26). However, this effect estimate was not statistically significant due to very serious imprecision and there was a high risk of selection bias. Moreover, this study (N=84) found no evidence for clinically or statistically significant effects of home visits on preventing poor maternal contact with primary and/or secondary care when an ITT analysis approach was used (p=0.60) (Table 65).

There was single study (N=131) evidence for a moderate benefit of home visits on preventing infant admissions to hospital (p=0.31) for women with psychosocial risk factors and (family) history of mental health problems (Table 65). However, confidence in this effect estimate was very low due to very serious imprecision (the event rate does not meet the rule-of-thumb threshold for optimal information size [Events<300] and the 95% CI includes no effect and measures of appreciable benefit and harm) and high risk of selective reporting bias. This same study found no evidence for a clinically or statistically significant effect of home visits on reducing infant length of stay in hospital (p=0.37).

Table 65: Summary of findings table for preventative effects of home visits compared with treatment as usual on service utilisation for women with identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Service utilisation: Home visits versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Maternal contact with primary and/or secondary care Post-treatment – ITT analysis (at-risk populations) Linkage with primary care (Has a regular personal doctor at year 2) Follow-up: mean 117 weeks	Study population 375 per 1000	431 per 1000 (255 to 731)	RR 1.15 (0.68 to 1.95)	84 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	
	Moderate 375 per 1000	431 per 1000 (255 to 731)				
Maternal contact with primary and/or secondary care Post-treatment – available case analysis (at-risk populations) Linkage with primary care (Has a regular personal doctor at year 2) Follow-up: mean 117 weeks	Study population 469 per 1000	614 per 1000 (384 to 975)	RR 1.31 (0.82 to 2.08)	63 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	
	Moderate 469 per 1000	614 per 1000 (385 to 976)				
Infant admissions to hospital Mid-treatment (at 6 months) – ITT analysis (at-risk populations) Infant service use: Admissions to hospital since birth Follow-up: mean 52 weeks	Study population 127 per 1000	74 per 1000 (25 to 213)	RR 0.58 (0.2 to 1.68)	131 (1 study)	⊕⊖⊖⊖ very low[2,3,4]	
	Moderate 127 per 1000	74 per 1000 (25 to 213)				
Infant length of stay in hospital Mid-treatment (at 6		The mean infant length of stay in hospital mid-treatment (at 6 months)		131 (1 study)	⊕⊖⊖⊖ very low[3,4,5]	SMD -0.16 (-0.5 to 0.19)

months) – ITT analysis (at-risk populations) Infant service use: Median days stayed in hospital Follow-up: mean 52 weeks	– ITT analysis (at-risk populations) in the intervention groups was 0.16 standard deviations lower (0.5 lower to 0.19 higher)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] High risk of selection bias due to unclear randomisation method and allocation concealment and statistically significant group difference at baseline (intervention group scored higher on measure of parenting attitudes and beliefs)
[2] Total number of events is less than 300 (a threshold rule-of-thumb)
[3] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
[4] Paper omits data
[5] Total population size is less than 400 (a threshold rule-of-thumb)

7.4.12 Clinical evidence for preventative effects on experience of care for women with identified risk factors (by intervention)

Summary of findings can be found in the tables presented in this section. The full GRADE evidence profiles and associated forest plots can be found in Appendix 22 and Appendix 19, respectively.

Experience of care: non-mental health-focused education and support versus treatment as usual or enhanced treatment as usual

A single study (N=141-162) found no evidence for clinically or statistically significant effects of non-mental health-focused education and support group and home visits relative to treatment as usual on preventing maternal dissatisfaction with care (p=0.09-0.15) for women with an uncomplicated twin pregnancy (Table 66).

Table 66: Summary of findings table for effects of non-mental health-focused education and support compared with treatment as usual or enhanced treatment as usual on preventing poor experience of care for women with identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk / Corresponding risk Control / Experience of care: Non-mental health-focused education and support versus TAU or Enhanced TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Maternal dissatisfaction with care Post-treatment – ITT analysis (at-risk populations) Self-report Follow-up: mean 6 weeks	Study population 634 per 1000 / 501 per 1000 (380 to 660) Moderate 634 per 1000 / 501 per 1000 (380 to 659)	RR 0.79 (0.6 to 1.04)	162 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	
Maternal dissatisfaction with care Post-treatment – available case analysis (at-risk populations) Self-report Follow-up: mean 6 weeks	Study population 565 per 1000 / 447 per 1000 (317 to 616) Moderate 565 per 1000 / 446 per 1000 (316 to 616)	RR 0.79 (0.56 to 1.09)	141 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total number of events is less than 300 (a threshold rule-of-thumb)
[2] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
[3] Paper omits data

7.4.13 Clinical evidence for preventative effects on poor retention in services and treatment unacceptability for women with identified risk factors (by intervention)

Summary of findings can be found in the tables presented in this section. The full GRADE evidence profiles and associated forest plots can be found in Appendix 22 and Appendix 19, respectively.

Retention in services and treatment acceptability (using attrition as a proxy measure): post-miscarriage self-help versus treatment as usual

A single study (N=228) found no evidence for clinically or statistically significant effects of post-miscarriage self-help on attrition (p=0.59) (Table 67).

Table 67: Summary of findings table for effects of post-miscarriage self-help compared with treatment as usual on preventing poor retention in services or treatment unacceptability for women with identified risk factors

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Attrition: post-miscarriage self-help versus TAU				
Drop-out Incomplete data at endpoint Follow-up: mean 5 weeks	Study population		RR 1.21 (0.61 to 2.4)	228 (1 study)	⊕⊖⊖⊖ very low1,2,3	
	115 per 1000	139 per 1000 (70 to 276)				
	Moderate					
	115 per 1000	139 per 1000 (70 to 276)				

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 High risk of selection bias due to unclear allocation concealment and statistically significant group differences at baseline
2 Total number of events is less than 300 (a threshold rule-of-thumb)
3 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Retention in services and treatment acceptability (using attrition as a proxy measure): social support versus treatment as usual

A single study (N=117) found evidence for a moderate harm associated with peer-mediated support (including one-to-one befriending and psychoeducational group meetings) with higher attrition in the intervention group relative to treatment as usual (p=0.15). However, this effect estimate was not statistically significant due to very serious imprecision (Table 68).

Table 68: Summary of findings table for effects of social support compared with treatment as usual on preventing poor retention in services or treatment unacceptability for women with identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	 Corresponding risk Attrition: Social support versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Drop-out Incomplete data at endpoint Follow-up: mean 12 weeks	Study population		RR 1.36 (0.89 to 2.06)	117 (1 study)	⊕⊕⊖⊖ low1,2	
	375 per 1000	510 per 1000 (334 to 772)				
	Moderate					
	375 per 1000	510 per 1000 (334 to 772)				

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total number of events is less than 300 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Retention in services and treatment acceptability (using attrition as a proxy measure): Psychologically (CBT/IPT)-informed psychoeducation versus treatment as usual or enhanced treatment as usual

There was evidence from three studies (N=360) for a moderate harm associated with CBT- or IPT-informed psychoeducation (p=0.42) with higher attrition in the intervention group relative to treatment as usual or enhanced treatment as usual

(non-mental health-focused education and support [booklet]). However, this effect was not statistically significant due to very serious imprecision (Table 69).

Table 69: Summary of findings table for effects of psychologically (CBT/IPT)-informed psychoeducation compared with treatment as usual or enhanced treatment as usual on preventing poor retention in services or treatment unacceptability for women with identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	 Corresponding risk Attrition: Psychologically (CBT/IPT)-informed psychoeducation versus TAU or Enhanced TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Drop-out Incomplete data at endpoint Follow-up: 26-27 weeks	Study population		RR 1.63 (0.5 to 5.28)	360 (3 studies)	⊕⊕⊖⊖ low1,2	
	67 per 1000	109 per 1000 (34 to 354)				
	Moderate					
	94 per 1000	153 per 1000 (47 to 496)				

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total number of events is less than 300 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Retention in services and treatment acceptability (using attrition as a proxy measure): Psychoeducational booklet versus treatment as usual or enhanced treatment as usual

A single study (N=600) found no evidence for clinically or statistically significant effects of a psychoeducational booklet relative to treatment as usual on attrition (p=0.23) for women with psychosocial risk factors and (family) history of mental health problems (Table 70).

Table 70: Summary of findings table for effects of psychoeducational booklet compared with treatment as usual or enhanced treatment as usual on preventing poor retention in services or treatment unacceptability for women with identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	 Corresponding risk Attrition: Psychoeducational booklet versus TAU or Enhanced TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Drop-out Incomplete data at endpoint	Study population		RR 0.88 (0.72 to 1.08)	600 (1 study)	⊕⊖⊖⊖ very low1,2,3	
	405 per 1000	357 per 1000 (292 to 438)				
	Moderate					
	405 per 1000	356 per 1000 (292 to 437)				

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 High risk of selection bias due to statistically significant group differences at baseline
2 Total number of events is less than 300 (a threshold rule-of-thumb)
3 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Retention in services and treatment acceptability (using attrition as a proxy measure): non-mental health-focused education and support versus treatment as usual or enhanced treatment as usual

There was evidence from three studies (N=584) for a moderate benefit of non-mental health focused education and support on preventing poor retention in services or treatment unacceptability (using attrition as a proxy measure) for women with a range of identified risk factors (p=0.06). However, confidence in this effect estimate is very low due to a high risk of selection bias (statistically significant group difference at baseline) and very serious imprecision (threshold rule-of-thumb for

optimal information size is not met and the 95% CI includes both no effect and measure of appreciable benefit) (Table 71).

Table 71: Summary of findings table for effects of non-mental health-focused education and support compared with treatment as usual or enhanced treatment as usual on preventing poor retention in services or treatment unacceptability for women with identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	 Corresponding risk Attrition: Non-mental health-focused education and support versus TAU or Enhanced TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Drop-out Incomplete data at endpoint Follow-up: 6-28 weeks	Study population		RR 0.72 (0.5 to 1.02)	584 (3 studies)	⊕⊖⊖⊖ very low1,2,3	
	209 per 1000	150 per 1000 (104 to 213)				
	Moderate					
	207 per 1000	149 per 1000 (104 to 211)				

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 High risk of selection bias due to a statistically significant group difference at baseline
2 Total number of events is less than 300 (a threshold rule-of-thumb)
3 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Retention in services and treatment acceptability (using attrition as a proxy measure): home visits versus treatment as usual

Two studies (N=215) found no evidence for clinically or statistically significant effects of home visits relative to treatment as usual on attrition (p=0.54; Table 72).

Table 72: Summary of findings table for effects of home visits compared with treatment as usual on preventing poor retention in services or treatment unacceptability for women with identified risk factors

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Attrition: Home visits versus TAU				
Drop-out Incomplete data at endpoint Follow-up: 78-117 weeks	Study population		**RR 1.23** (0.64 to 2.37)	215 (2 studies)	⊕⊖⊖⊖ **very low**[1,2,3]	
	126 per 1000	155 per 1000 (81 to 299)				
	Moderate					
	140 per 1000	172 per 1000 (90 to 332)				

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 High risk of selection bias due to unclear randomisation method and statistically significant group difference at baseline
2 Total number of events is less than 300 (a threshold rule-of-thumb)
3 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Retention in services and treatment acceptability (using attrition as a proxy measure): post-delivery discussion versus enhanced treatment as usual

There was single study (N=1,041) evidence for a moderate effect of a midwife-led post-delivery discussion relative to enhanced treatment as usual (non-mental health-focused information [booklet]) on preventing poor retention in services and treatment unacceptability (using attrition as a proxy) for women who had had an operative delivery (p=0.09). However, this effect was not statistically significant due to very serious imprecision (Table 73).

Table 73: Summary of findings table for effects of post-delivery discussion compared with enhanced treatment as usual on preventing poor retention in services or treatment unacceptability for women with identified risk factors

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Attrition: post-delivery discussion versus Enhanced TAU				
Drop-out Incomplete data at endpoint Follow-up: mean 26 weeks	Study population		RR 0.75 (0.54 to 1.04)	1,041 (1 study)	⊕⊕⊖⊖ low1,2	
	136 per 1000	102 per 1000 (74 to 142)				
	Moderate					
	136 per 1000	102 per 1000 (73 to 141)				

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total number of events is less than 300 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Retention in services and treatment acceptability (using attrition as a proxy measure): mother-infant relationship interventions versus treatment as usual

Four studies (N=772) found no evidence for clinically or statistically significant effects of mother–infant relationship interventions relative to treatment as usual on attrition (p=0.79) for women with psychosocial risk factors or who had had a preterm delivery and/or low birthweight baby (Table 74).

Table 74: Summary of findings table for effects of mother–infant relationship interventions compared with treatment as usual on preventing poor retention in services or treatment unacceptability for women with identified risk factors

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Attrition: Mother-infant relationship interventions versus TAU				
Drop-out Incomplete data at endpoint Follow-up: 15-26 weeks	Study population		RR 1.04 (0.76 to 1.43)	772 (4 studies)	⊕⊕⊖⊖ low1,2	
	201 per 1000	209 per 1000 (152 to 287)				
	Moderate					
	168 per 1000	175 per 1000 (128 to 240)				

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total number of events is less than 300 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

7.4.14 Clinical evidence for preventative effects on infant physical health problems where mothers have identified risk factors (by intervention)

Summary of findings can be found in the tables presented in this section. The full GRADE evidence profiles and associated forest plots can be found in Appendix 22 and Appendix 19, respectively.

Infant physical health: home visits versus treatment as usual

A single study (N=131) found low quality evidence for a large harm associated with home visits for women with psychosocial risk factors and (family) history of mental health problems, with a larger number of infants found with congenital malformations/disabilities (measured at 6 months) in the intervention relative to the

control group (p=0.11). However, this effect was not statistically significant due to very serious imprecision (the threshold rule-of-thumb for the optimal information size, that is 300 events, was not met and the 95% CI includes no effect and measures of both appreciable benefit and appreciable harm) (Table 75).

Another single study (N=79) found very low quality evidence for a moderate benefit of home visits for adolescent mothers with psychosocial risk factors in preventing infants being underweight (p=0.43). However, this effect was not statistically significant due to very serious imprecision and there are risk of bias concerns due to unclear selection and detection bias (Table 75). The same study (N=79-87) found no evidence for clinically or statistically significant effects of home visits on increasing the number of infants of normal weight (p=0.72) or preventing infants from being overweight (p=0.86) or preventing the incidence of severe diarrhoea for infants (p=0.81).

Table 75: Summary of findings table for effects of home visits compared with treatment as usual on preventing poor physical health in infants where mothers have identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Infant physical health: Home visits versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Congenital malformations (measured at 6 months) – available case analysis (at-risk populations) Number of infants with a disability Follow-up: mean 52 weeks	Study population		RR 5.56 (0.69 to 44.9)	131 (1 study)	⊕⊕⊖⊖ low1,2	
	16 per 1000	88 per 1000 (11 to 713)				
	Moderate					
	16 per 1000	89 per 1000 (11 to 718)				
Normal weight Post-treatment – available case analysis (at-risk populations) Number of infants of a normal weight	Study population		RR 1.09 (0.68 to 1.75)	79 (1 study)	⊕⊖⊖⊖ very low1,2,3	
	447 per 1000	488 per 1000 (304 to 783)				
	Moderate					
	447 per 1000	487 per 1000 (304 to 782)				
Underweight Post-treatment – available case analysis (at-risk populations)	Study population		RR 0.62 (0.19 to 2.02)	79 (1 study)	⊕⊖⊖⊖ very low1,2,3	
	158 per 1000	98 per 1000 (30 to 319)				
	Moderate					

Number of infants who are underweight	158 per 1000	98 per 1000 (30 to 319)			
Overweight Post-treatment – available case analysis (at-risk populations) Number of infants who are overweight	Study population 395 per 1000	414 per 1000 (241 to 711)	RR 1.05 (0.61 to 1.8)	79 (1 study)	⊕⊖⊖⊖ very low1,2,3
	Moderate 395 per 1000	415 per 1000 (241 to 711)			
Incidence of severe diarrhoea Post-treatment – available case analysis (at-risk populations) Infant illness: Severe diarrhoea (without dehydration)	Study population 95 per 1000	111 per 1000 (32 to 386)	RR 1.17 (0.34 to 4.05)	87 (1 study)	⊕⊖⊖⊖ very low1,2,3
	Moderate 95 per 1000	111 per 1000 (32 to 385)			

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total number of events is less than 300 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
3 Unclear risk of selection bias due to insufficient detail reported with regards to randomisation method and allocation concealment and unclear risk of detection bias as blinding of outcome assessor not reported

7.4.15 Clinical evidence for preventative effects on infant regulatory problems where mothers have identified risk factors (by intervention)

Summary of findings can be found in the tables presented in this section. The full GRADE evidence profiles and associated forest plots can be found in Appendix 22 and Appendix 19, respectively.

Infant regulatory problems: mother-infant relationship interventions versus treatment as usual

A single study (N=63) found evidence for moderate to very large effects of a mother–infant relationship intervention relative to treatment as usual for mothers who had had a preterm delivery on preventing infant colic (at post-treatment [p <0.0001] and short-term follow-up [p <0.00001]), infant sleep problems (at post-treatment [p <0.00001] and short-term follow-up [p=0.02]), and infant excessive

crying (at post-treatment [p <0.0001] but not at short-term follow-up [p=0.09]). However, confidence in these effect estimates is very low to very serious imprecision (very small sample size) and a high risk of selective reporting bias (Table 76).

Table 76: Summary of findings table for effects of mother–infant relationship interventions compared with treatment as usual on preventing regulatory problems in infants where mothers have identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Infant regulatory problems: Mother-infant relationship interventions versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Infant colic mean scores post-treatment – available case analysis (at-risk populations) Short Temperament Scale for Infants (STSI): Colic Follow-up: mean 15 weeks		The mean infant colic mean scores post-treatment – available case analysis (at-risk populations) in the intervention groups was 1.08 standard deviations lower (1.61 to 0.55 lower)		63 (1 study)	⊕⊖⊖⊖ very low1,2	SMD -1.08 (-1.61 to -0.55)
Infant sleep problems mean score Post-treatment – available case analysis (at-risk populations) STSI: Sleep problems Follow-up: mean 15 weeks		The mean infant sleep problems mean score post-treatment – available case analysis (at-risk populations) in the intervention groups was 5.27 standard deviations lower (6.34 to 4.2 lower)		63 (1 study)	⊕⊖⊖⊖ very low1,2	SMD -5.27 (-6.34 to -4.2)
Infant excessive crying mean scores post-treatment – available case analysis (at-risk populations) STSI: Excessive crying Follow-up: mean 15 weeks		The mean infant excessive crying mean scores post-treatment – available case analysis (at-risk populations) in the intervention groups was 1.13 standard deviations lower (1.67 to 0.6 lower)		63 (1 study)	⊕⊖⊖⊖ very low1,2	SMD -1.13 (-1.67 to -0.6)

Infant colic mean scores Short follow-up (9-16 weeks post-intervention) – available case analysis (at-risk populations) STSI: Colic Follow-up: mean 28 weeks	The mean infant colic mean scores short follow-up (9-16 weeks post-intervention) – available case analysis (at-risk populations) in the intervention groups was 1.72 standard deviations lower (2.31 to 1.14 lower)	63 (1 study)	⊕⊖⊖⊖ very low1,2	SMD -1.72 (-2.31 to -1.14)
Infant sleep problems mean score Short follow-up (9-16 weeks post-intervention) – available case analysis (at-risk populations) STSI: Sleep problems Follow-up: mean 28 weeks	The mean infant sleep problems mean score short follow-up (9-16 weeks post-intervention) – available case analysis (at-risk populations) in the intervention groups was 0.6 standard deviations lower (1.1 to 0.09 lower)	63 (1 study)	⊕⊖⊖⊖ very low1,2	SMD -0.6 (-1.1 to -0.09)
Infant excessive crying mean scores Short follow-up (9-16 weeks post-intervention) – available case analysis (at-risk populations) STSI: Excessive crying Follow-up: mean 28 weeks	The mean infant excessive crying mean scores short follow-up (9-16 weeks post-intervention) – available case analysis (at-risk populations) in the intervention groups was 0.43 standard deviations lower (0.93 lower to 0.07 higher)	63 (1 study)	⊕⊖⊖⊖ very low1,2,3	SMD -0.43 (-0.93 to 0.07)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total population size is less than 400 (a threshold rule-of-thumb)
2 Paper omits data
3 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

7.4.16 Clinical evidence for preventative effects on infant physical development problems where mothers have identified risk factors (by intervention)

Summary of findings can be found in the tables presented in this section. The full GRADE evidence profiles and associated forest plots can be found in Appendix 22 and Appendix 19, respectively.

Infant physical development: home visits versus treatment as usual

Two studies (N=194) found evidence for a moderate effect of home visits, for adolescent mothers with psychosocial risk factors or mothers who had had a preterm delivery, for preventing delayed or impaired motor development when an available case analysis approach was used (p=0.54). However, confidence in this effect estimate was very low due to risk of bias concerns (statistically significant group difference at baseline), very serious imprecision (the rule-of-thumb threshold for optimal information size was not met [Events<300] and the 95% CI includes no effect and measures of both appreciable benefit and appreciable harm) and there was a high risk of selective reporting bias (Table 77). Moreover, a single study (N=96-120) found no evidence for clinically or statistically significant effects of home visits on preventing delayed or impaired motor development at long-term follow-up when an available case analysis approach was used (p=0.71) or at post-treatment (p=0.74) or long-term follow-up (p=0.82) when an ITT analysis approach was used, and up to two studies (N=96-194) found no evidence for clinically or statistically significant effects of home visits on preventing poor motor development mean scores at post-treatment (p=0.87) or long-term follow-up (p=0.88).

Table 77: Summary of findings table for effects of home visits compared with treatment as usual on preventing physical development problems in infants where mothers have identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Infant physical development: Home visits versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Infant motor development (delayed or impaired) Post-treatment – ITT analysis (at-risk populations) Bayley Scales of Infant Development-	Study population		RR 0.86 (0.36 to 2.08)	120 (1 study)	⊕⊖⊖⊖ very low1,2,3,4	
	153 per 1000	131 per 1000 (55 to 317)				
	Moderate					
	153 per 1000	132 per 1000 (55 to 318)				

Motor (scores<70) Follow-up: mean 104 weeks						
Infant motor development (delayed or impaired) Post-treatment – available case analysis (at-risk populations) Psychomotor Development Scale-General Development (at risk or delayed) or Bayley Scales of Infant Development-Motor (scores<70) Follow-up: mean 104 weeks	Study population 84 per 1000 Moderate 75 per 1000	61 per 1000 (23 to 168) 55 per 1000 (20 to 150)	RR 0.73 (0.27 to 2)	194 (2 studies)	⨁⊖⊖⊖ very low1,2,3,4	
Infant motor development mean scores post-treatment – available case analysis (at-risk populations) Psychomotor Development Scale-General Development or Bayley Scales of Infant Development-Motor Follow-up: mean 104 weeks		The mean infant motor development mean scores post-treatment – available case analysis (at-risk populations) in the intervention groups was 0.02 standard deviations higher (0.26 lower to 0.3 higher)		194 (2 studies)	⨁⊖⊖⊖ very low1,4,5	SMD 0.02 (-0.26 to 0.3)
Infant motor development (delayed or impaired) Long follow-up (25-103 weeks post-intervention) – ITT analysis (at-risk populations) Movement Assessment Battery for Children: Total motor problems (scores=<15th percentile) Follow-up: mean 208 weeks	Study population 373 per 1000 Moderate 373 per 1000	395 per 1000 (250 to 619) 395 per 1000 (250 to 619)	RR 1.06 (0.67 to 1.66)	120 (1 study)	⨁⊖⊖⊖ very low1,2,3,4	
	Study population					

Infant motor development (delayed or impaired) Long follow-up (25-103 weeks post-intervention) – available case analysis (at-risk populations) Movement Assessment Battery for Children: Total motor problems (scores=<15th percentile) Follow-up: mean 208 weeks	213 per 1000	245 per 1000 (117 to 513)	RR 1.15 (0.55 to 2.41)	96 (1 study)	⊕⊖⊖⊖ very low1,2,3,4	
	Moderate					
	213 per 1000	245 per 1000 (117 to 513)				
Infant motor development mean scores Long follow-up (25-103 weeks post-intervention) – available case analysis (at-risk populations) Movement Assessment Battery for Children: Total motor problems Follow-up: mean 208 weeks		The mean infant motor development mean scores long follow-up (25-103 weeks post-intervention) – available case analysis (at-risk populations) in the intervention groups was 0.03 standard deviations lower (0.43 lower to 0.37 higher)		96 (1 study)	⊕⊖⊖⊖ very low1,4,5	SMD -0.03 (-0.43 to 0.37)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 High risk of selection bias due to statistcially significant baseline difference between groups with twice the number of participants showing depression symptomatology (EPDS=>13) in the control group (N=10/17%) relative to the intervention group (N=5/8%)
2 Total number of events is less than 300 (a threshold rule-of-thumb)
3 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
4 Paper omits data
5 Total population size is less than 400 (a threshold rule-of-thumb)

7.4.17 Clinical evidence for preventative effects on infant cognitive development problems where mothers have identified risk factors (by intervention)

Summary of findings can be found in the tables presented in this section. The full GRADE evidence profiles and associated forest plots can be found in Appendix 22 and Appendix 19, respectively.

Infant cognitive development: home visits versus treatment as usual

A single study (N=101) found evidence for a large harm associated with home visits for infants of women who had had a preterm delivery with a greater number of infants in the intervention group relative to treatment as usual showing nonverbal development impairment at post-treatment when an available case analysis approach was used (p=0.19). However, confidence in this effect estimate was very low due to high risk of selection and selective reporting bias and very serious imprecision, and the effect estimate for this outcome measure was not statistically or clinically significant when an ITT analysis approach was used (N=120; p=0.48). This same study (N=104) also found evidence for a large benefit associated with home visits on preventing infant verbal development impairment at long-term follow-up when an available case analysis was used (p=0.15), however, again confidence in this effect estimate was very low due to risk of bias concerns and very serious imprecision and the effect estimate was not clinically or statistically significant when an ITT analysis approach was used (p=0.46), or at post-treatment using either analysis approach (N=111-120; p=0.89-0.91). This study (N=99-120) found no evidence for clinically or statistically significant effects of home visits for preventing infant: cognitive development impairment (at post-treatment [p=0.74-0.94] or long-term follow [p=0.77-0.82]); poor cognitive development mean scores (at post-treatment [p=0.16] or long-term follow-up [p=0.65]); poor verbal development mean scores (at post-treatment [p=0.63] or long-term follow-up [p=0.15]); poor nonverbal development mean scores (at first measurement [p=0.30]); spatial reasoning impairment (at first measurement [p=0.94-0.96]); poor spatial reasoning mean scores (at first measurement [p=0.49]) (Table 78).

Table 78: Summary of findings table for effects of home visits compared with treatment as usual on preventing cognitive development problems in infants where mothers have identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Infant cognitive development: Home visits versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Infant cognitive development (impairment) Post-treatment – ITT analysis (at-risk populations) Bayley Scales of Infant Development-Cognitive (scores <70) Follow-up: mean 104 weeks	Study population		**RR 0.97** (0.41 to 2.27)	120 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3,4]	
	153 per 1000	148 per 1000 (63 to 346)				
	Moderate					
	153 per 1000	148 per 1000 (63 to 347)				
Infant cognitive development (impairment) Post-treatment – available case analysis (at-risk populations) Bayley Scales of Infant Development-Cognitive (scores <70) Follow-up: mean 104 weeks	Study population		**RR 0.84** (0.3 to 2.35)	115 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3,4]	
	123 per 1000	103 per 1000 (37 to 289)				
	Moderate					
	123 per 1000	103 per 1000 (37 to 289)				
Infant cognitive development mean scores post-treatment – available case analysis (at-risk populations) Bayley Scales of Infant Development – Cognitive Follow-up: mean 104 weeks		The mean infant cognitive development mean scores post-treatment – available case analysis (at-risk populations) in the intervention groups was 0.27 standard deviations higher (0.1 lower to 0.63 higher)		115 (1 study)	⊕⊖⊖⊖ **very low**[1,3,4,5]	SMD 0.27 (-0.1 to 0.63)
Infant verbal development (impairment) Post-treatment – ITT	Study population		**RR 1.04** (0.55 to 1.95)	120 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3,4]	
	237 per 1000	247 per 1000 (131 to 463)				
	Moderate					

analysis (at-risk populations) Bayley Scales of Infant Development-Language (scores <70) Follow-up: mean 104 weeks	237 per 1000	246 per 1000 (130 to 462)				
Infant verbal development (impairment) Post-treatment – available case analysis (at-risk populations) Bayley Scales of Infant Development-Language (scores <70) Follow-up: mean 104 weeks	Study population 204 per 1000 Moderate 204 per 1000	194 per 1000 (92 to 407) 194 per 1000 (92 to 408)	**RR 0.95** (0.45 to 2)	111 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3,4]	
Infant verbal development mean scores post-treatment – available case analysis (at-risk populations) Bayley Scales of Infant Development-Language Follow-up: mean 104 weeks		The mean infant verbal development mean scores post-treatment – available case analysis (at-risk populations) in the intervention groups was 0.09 standard deviations lower (0.47 lower to 0.28 higher)		111 (1 study)	⊕⊖⊖⊖ **very low**[1,4,5]	SMD -0.09 (-0.47 to 0.28)
Infant nonverbal development (impairment) Post-treatment – ITT analysis (at-risk populations) Differential Abilities Scale: Nonverbal Reasoning composite (scores >1 SD below test mean) Follow-up: mean 208 weeks	Study population 237 per 1000 Moderate 237 per 1000	294 per 1000 (161 to 539) 294 per 1000 (161 to 538)	**RR 1.24** (0.68 to 2.27)	120 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3,4]	
Infant nonverbal development (impairment) Post-treatment – available case analysis (at-risk populations) Differential Abilities Scale: Nonverbal Reasoning composite (scores >1	Study population 82 per 1000 Moderate 82 per 1000	173 per 1000 (57 to 526) 174 per 1000 (57 to 528)	**RR 2.12** (0.7 to 6.44)	101 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3,4]	

SD below test mean) Follow-up: mean 208 weeks					
Infant nonverbal development mean scores post-treatment – available case analysis (at-risk populations) Differential Abilities Scale: Nonverbal Reasoning composite Follow-up: mean 208 weeks		The mean infant nonverbal development mean scores post-treatment – available case analysis (at-risk populations) in the intervention groups was 0.2 standard deviations lower (0.6 lower to 0.19 higher)		101 (1 study)	⊕⊖⊖⊖ **very low**[1,3,4,5]	SMD -0.2 (-0.6 to 0.19)
Infant spatial reasoning development (impairment) Post-treatment – ITT analysis (at-risk populations) Differential Abilities Scale: Spatial Reasoning composite (scores >1 SD below test mean) Follow-up: mean 208 weeks	Study population 305 per 1000	311 per 1000 (183 to 534)	**RR 1.02** (0.6 to 1.75)	120 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3,4]	
	Moderate 305 per 1000	311 per 1000 (183 to 534)				
Infant spatial reasoning development (impairment) Post-treatment – available case analysis (at-risk populations) Differential Abilities Scale: Spatial Reasoning composite (scores >1 SD below test mean) Follow-up: mean 208 weeks	Study population 163 per 1000	160 per 1000 (65 to 392)	**RR 0.98** (0.4 to 2.4)	99 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3,4]	
	Moderate 163 per 1000	160 per 1000 (65 to 391)				
Infant spatial reasoning development mean scores post-treatment – available case analysis (at-risk populations) Differential Abilities Scale: Spatial Reasoning composite		The mean infant spatial reasoning development mean scores post-treatment – available case analysis (at-risk populations) in the intervention groups was 0.14 standard deviations higher (0.26 lower to 0.53 higher)		99 (1 study)	⊕⊖⊖⊖ **very low**[1,3,4,5]	SMD 0.14 (-0.26 to 0.53)

Follow-up: mean 208 weeks					
Infant cognitive development (impairment) Long Follow-up (25-103 weeks post-intervention) – ITT analysis (at-risk populations) Differential Abilities Scale: General Conceptual Ability (scores >1 SD below test mean) Follow-up: mean 208 weeks	Study population 271 per 1000	296 per 1000 (168 to 521) Moderate	**RR 1.09** (0.62 to 1.92)	120 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3,4]
	Moderate 271 per 1000	295 per 1000 (168 to 520)			
Infant cognitive development (impairment) Long Follow-up (25-103 weeks post-intervention) – available case analysis (at-risk populations) Differential Abilities Scale: General Conceptual Ability (scores >1 SD below test mean) Follow-up: mean 208 weeks	Study population 157 per 1000	173 per 1000 (72 to 414) Moderate	**RR 1.1** (0.46 to 2.64)	103 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3,4]
	Moderate 157 per 1000	173 per 1000 (72 to 414)			
Infant cognitive development mean scores Long Follow-up (25-103 weeks post-intervention) – available case analysis (at-risk populations) Differential Abilities Scale: General Conceptual Ability Follow-up: mean 208 weeks		The mean infant cognitive development mean scores long follow-up (25-103 weeks post-intervention) – available case analysis (at-risk populations) in the intervention groups was 0.09 standard deviations higher (0.3 lower to 0.48 higher)		103 (1 study)	⊕⊖⊖⊖ **very low**[1,4,5] SMD 0.09 (-0.3 to 0.48)
Infant verbal development (impairment) Long Follow-up (25-103 weeks post-intervention) – ITT analysis (at-risk populations)	Study population 271 per 1000	214 per 1000 (114 to 404) Moderate	**RR 0.79** (0.42 to 1.49)	120 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3,4]
	Moderate 271 per 1000	214 per 1000 (114 to 404)			

Differential Abilities Scale: Verbal composite (scores >1 SD below test mean) Follow-up: mean 208 weeks					
Infant verbal development (impairment) Long Follow-up (25-103 weeks post-intervention) – available case analysis (at-risk populations) Differential Abilities Scale: Verbal composite (scores >1 SD below test mean) Follow-up: mean 208 weeks	Study population 173 per 1000	76 per 1000 (26 to 234)	**RR 0.44** (0.15 to 1.35)	104 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3,4]
	Moderate 173 per 1000	76 per 1000 (26 to 234)			
Infant verbal development mean scores Long Follow-up (25-103 weeks post-intervention) – available case analysis (at-risk populations) Differential Abilities Scale: Verbal composite Follow-up: mean 208 weeks		The mean infant verbal development mean scores long follow-up (25-103 weeks post-intervention) – available case analysis (at-risk populations) in the intervention groups was 0.28 standard deviations higher (0.1 lower to 0.67 higher)		104 (1 study)	⊕⊖⊖⊖ **very low**[1,3,4,5] SMD 0.28 (-0.1 to 0.67)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] High risk of selection bias due to statistically significant baseline difference between groups with twice the number of participants showing depression symptomatology (EPDS ≥13) in the control group (N=10/17%) relative to the intervention group (N=5/8%)
[2] Total number of events is less than 300 (a threshold rule-of-thumb)
[3] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
[4] Paper omits data
[5] Total population size is less than 400 (a threshold rule-of-thumb)

7.4.18 Clinical evidence for preventative effects on infant emotional development problems where mothers have identified risk factors (by intervention)

Summary of findings can be found in the tables presented in this section. The full GRADE evidence profiles and associated forest plots can be found in Appendix 22 and Appendix 19, respectively.

Infant emotional development: home visits versus treatment as usual

There was single study (N=97-120) evidence for small to large effects of home visits for women who had had a preterm delivery on preventing infant adaptive behaviour impairment (p=0.07), poor adaptive behaviour mean scores (p=0.02), externalizing impairment (p=0.08), higher externalizing mean scores (p=0.03) or internalizing impairment (p=0.44) at post-treatment and higher internalizing mean scores at long-term follow-up (p=0.02) when an available case analysis approach was used (Table 79). However, the effect estimates for the same outcome measures were not clinically or statistically significant when an ITT analysis approach was adopted (p=0.37-0.73). Effects on overall emotional development (impairment on one or more domain [p=0.03-0.005]) and dysregulation impairment (p=0.03-0.09) were, however, either clinically significant or both clinically and statistically significant using either analysis approach. There was also evidence for a large effect on preventing higher dysregulation mean scores (p=0.0001). However, confidence in all these effect estimates was very low due to a high risk of selection and selective reporting bias and very serious imprecision. This study found no evidence for clinically or statistically significant effects on preventing: higher internalizing mean scores (p=0.45) at post-treatment; adaptive behaviour impairment (p=0.37-0.60); poorer adaptive behaviour mean scores (p=0.35) at long-term follow-up; higher externalizing mean scores at long-term follow-up (p=0.80); internalizing impairment at long-term follow-up (p=0.48-0.63). There was evidence for a moderate harm associated with home visits on externalizing impairment at long-term follow-up when an available case analysis approach was used (p=0.43) but not when an ITT approach was adopted (p=0.97).

Table 79: Summary of findings table for effects of home visits compared with treatment as usual on preventing emotional development problems in infants where mothers have identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Infant emotional development: Home visits versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Infant adaptive behaviour (impairment) Post-treatment – ITT analysis (at-risk populations) Infant Toddler Social and Emotional Assessment: Competence (mean scores ≤=10th percentile) Follow-up: mean 104 weeks	Study population 390 per 1000	312 per 1000 (191 to 511)	**RR 0.8** (0.49 to 1.31)	120 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3,4]	
	Moderate 390 per 1000	312 per 1000 (191 to 511)				
Infant adaptive behaviour (impairment) Post-treatment – available case analysis (at-risk populations) Infant Toddler Social and Emotional Assessment: Competence (mean scores ≤=10th percentile) Follow-up: mean 104 weeks	Study population 306 per 1000	147 per 1000 (64 to 324)	**RR 0.48** (0.21 to 1.06)	97 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3,4]	
	Moderate 306 per 1000	147 per 1000 (64 to 324)				
Infant adaptive behaviour mean scores post-treatment – available case analysis (at-risk populations) Infant Toddler Social and Emotional Assessment: Competence Follow-up: mean 104 weeks		The mean infant adaptive behaviour mean scores post-treatment – available case analysis (at-risk populations) in the intervention groups was 0.49 standard deviations higher (0.09 to 0.89 higher)		99 (1 study)	⊕⊖⊖⊖ **very low**[1,4,5]	SMD 0.49 (0.09 to 0.89)

Infant emotional development (impairment) Post-treatment – ITT analysis (at-risk populations) Infant Toddler Social and Emotional Assessment: Impairment ≥1 domain Follow-up: mean 104 weeks	Study population 559 per 1000	358 per 1000 (241 to 543)	**RR 0.64** (0.43 to 0.97)	120 (1 study)	⊕⊖⊖⊖ **very low**[1,2,4]
	Moderate 559 per 1000	358 per 1000 (240 to 542)			
Infant emotional development (impairment) Post-treatment – available case analysis (at-risk populations) Infant Toddler Social and Emotional Assessment: Impairment ≥1 domain Follow-up: mean 104 weeks	Study population 500 per 1000	210 per 1000 (110 to 385)	**RR 0.42** (0.22 to 0.77)	98 (1 study)	⊕⊖⊖⊖ **very low**[1,2,4]
	Moderate 500 per 1000	210 per 1000 (110 to 385)			
Infant externalizing (impairment) Post-treatment – ITT analysis (at-risk populations) Infant Toddler Social and Emotional Assessment: Externalizing (mean scores ≥90th percentile) Follow-up: mean 104 weeks	Study population 271 per 1000	231 per 1000 (122 to 428)	**RR 0.85** (0.45 to 1.58)	120 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3,4]
	Moderate 271 per 1000	230 per 1000 (122 to 428)			
Infant externalizing (impairment) Post-treatment – available case analysis (at-risk populations) Infant Toddler Social and Emotional Assessment: Externalizing (mean scores ≥90th percentile) Follow-up: mean 104 weeks	Study population 157 per 1000	41 per 1000 (9 to 184)	**RR 0.26** (0.06 to 1.17)	100 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3,4]
	Moderate 157 per 1000	41 per 1000 (9 to 184)			
Infant externalizing mean scores post-		The mean infant externalizing mean		100 (1 study)	⊕⊖⊖⊖ **very low**[1,4,5]	SMD -0.43 (-0.83 to -0.03)

treatment – available case analysis (at-risk populations) Infant Toddler Social and Emotional Assessment: Externalizing Follow-up: mean 104 weeks		scores post-treatment – available case analysis (at-risk populations) in the intervention groups was 0.43 standard deviations lower (0.83 to 0.03 lower)			
Infant internalizing (impairment) Post-treatment – ITT analysis (at-risk populations) Infant Toddler Social and Emotional Assessment: Internalizing (mean scores ≥90th percentile) Follow-up: mean 104 weeks	Study population 203 per 1000	230 per 1000 (116 to 454)	**RR 1.13** (0.57 to 2.23)	120 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3,4]
	Moderate 203 per 1000	229 per 1000 (116 to 453)			
Infant internalizing (impairment) Post-treatment – available case analysis (at-risk populations) Infant Toddler Social and Emotional Assessment: Internalizing (mean scores ≥90th percentile) Follow-up: mean 104 weeks	Study population 78 per 1000	41 per 1000 (8 to 213)	**RR 0.52** (0.1 to 2.71)	100 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3,4]
	Moderate 78 per 1000	41 per 1000 (8 to 211)			
Infant internalizing mean scores post-treatment – available case analysis (at-risk populations) Infant Toddler Social and Emotional Assessment: Internalizing Follow-up: mean 104 weeks		The mean infant internalizing mean scores post-treatment – available case analysis (at-risk populations) in the intervention groups was 0.15 standard deviations lower (0.54 lower to 0.24 higher)		100 (1 study)	⊕⊖⊖⊖ **very low**[1,3,4,5] SMD -0.15 (-0.54 to 0.24)
Infant dysregulation (impairment) Post-treatment – ITT analysis (at-risk populations) Infant Toddler Social and Emotional Assessment:	Study population 339 per 1000	197 per 1000 (105 to 366)	**RR 0.58** (0.31 to 1.08)	120 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3,4]
	Moderate 339 per 1000	197 per 1000 (105 to 366)			

Dysregulation (mean scores ≥90th percentile) Follow-up: mean 104 weeks						
Infant dysregulation (impairment) Post-treatment – available case analysis (at-risk populations) Infant Toddler Social and Emotional Assessment: Dysregulation (mean scores ≥90th percentile) Follow-up: mean 104 weeks	Study population 235 per 1000	9 per 1000 (0 to 160)	**RR 0.04** (0 to 0.68)	100 (1 study)	⊕⊖⊖⊖ **very low**[1,2,4]	
	Moderate 235 per 1000	9 per 1000 (0 to 160)				
Infant dysregulation mean scores post-treatment – available case analysis (at-risk populations) Infant Toddler Social and Emotional Assessment: Dysregulation Follow-up: mean 104 weeks		The mean infant dysregulation mean scores post-treatment – available case analysis (at-risk populations) in the intervention groups was 0.8 standard deviations lower (1.21 to 0.39 lower)		100 (1 study)	⊕⊖⊖⊖ **very low**[1,4,5]	SMD -0.8 (-1.21 to -0.39)
Infant adaptive behaviour (impairment) Long Follow-up (25-103 weeks post-intervention) – ITT analysis (at-risk populations) Behavioral Assessment Screener for Children: Adaptive skills (scores>1 SD below test mean) Follow-up: mean 208 weeks	Study population 441 per 1000	361 per 1000 (234 to 560)	**RR 0.82** (0.53 to 1.27)	120 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3,4]	
	Moderate 441 per 1000	362 per 1000 (234 to 560)				
Infant adaptive behaviour (impairment) Long Follow-up (25-103 weeks post-intervention) – available case analysis (at-risk populations)	Study population 214 per 1000	169 per 1000 (73 to 401)	**RR 0.79** (0.34 to 1.87)	89 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3,4]	
	Moderate 214 per 1000	169 per 1000 (73 to 400)				

Behavioral Assessment Screener for Children: Adaptive skills (scores>1 SD below test mean) Follow-up: mean 208 weeks					
Infant adaptive behaviour mean scores Long Follow-up (25-103 weeks post-intervention) – available case analysis (at-risk populations) Behavioral Assessment Screener for Children: Adaptive skills Follow-up: mean 208 weeks	The mean infant adaptive behaviour mean scores long follow-up (25-103 weeks post-intervention) – available case analysis (at-risk populations) in the intervention groups was 0.2 standard deviations higher (0.22 lower to 0.62 higher)		89 (1 study)	⊕⊖⊖⊖ **very low**[1,3,4,5]	SMD 0.2 (-0.22 to 0.62)
Infant externalizing (impairment) Long Follow-up (25-103 weeks post-intervention) – ITT analysis (at-risk populations) Behavioral Assessment Screener for Children: Externalizing (scores >1 SD above test mean) Follow-up: mean 208 weeks	Study population 407 per 1000 — 411 per 1000 (264 to 631) Moderate 407 per 1000 — 411 per 1000 (265 to 631)	**RR 1.01** (0.65 to 1.55)	120 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3,4]	
Infant externalizing (impairment) Long Follow-up (25-103 weeks post-intervention) – available case analysis (at-risk populations) Behavioral Assessment Screener for Children: Externalizing (scores >1 SD above test mean) Follow-up: mean 208 weeks	Study population 167 per 1000 — 233 per 1000 (100 to 548) Moderate 167 per 1000 — 234 per 1000 (100 to 549)	**RR 1.4** (0.6 to 3.29)	89 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3,4]	

Infant externalizing mean scores Long Follow-up (25-103 weeks post-intervention) – available case analysis (at-risk populations) Behavioral Assessment Screener for Children: Externalizing Follow-up: mean 208 weeks		The mean infant externalizing mean scores long follow-up (25-103 weeks post-intervention) – available case analysis (at-risk populations) in the intervention groups was 0.05 standard deviations lower (0.47 lower to 0.36 higher)		89 (1 study)	⊕⊖⊖⊖ **very low**[1,4,5]	SMD -0.05 (-0.47 to 0.36)
Infant internalizing (impairment) Long Follow-up (25-103 weeks post-intervention) – ITT analysis (at-risk populations) Behavioral Assessment Screener for Children: Internalizing (score >1 SD above test mean) Follow-up: mean 208 weeks	Study population 407 per 1000 Moderate 407 per 1000	346 per 1000 (216 to 549) 346 per 1000 (216 to 549)	**RR 0.85** (0.53 to 1.35)	120 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3,4]	
Infant internalizing (impairment) Long Follow-up (25-103 weeks post-intervention) – available case analysis (at-risk populations) Behavioral Assessment Screener for Children: Internalizing (scores >1 SD above test mean) Follow-up: mean 208 weeks	Study population 167 per 1000 Moderate 167 per 1000	130 per 1000 (48 to 357) 130 per 1000 (48 to 357)	**RR 0.78** (0.29 to 2.14)	88 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3,4]	
Infant internalizing mean scores Long Follow-up (25-103 weeks post-intervention) – available case analysis (at-risk populations) Behavioral Assessment Screener		The mean infant internalizing mean scores long follow-up (25-103 weeks post-intervention) – available case analysis (at-risk populations) in the intervention groups was 0.5 standard deviations		88 (1 study)	⊕⊖⊖⊖ **very low**[1,4,5]	SMD -0.5 (-0.93 to -0.08)

for Children: Internalizing Follow-up: mean 208 weeks	lower (0.93 to 0.08 lower)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] High risk of selection bias due to statistically significant baseline difference between groups with twice the number of participants showing depression symptomatology (EPDS ≥13) in the control group (N=10/17%) relative to the intervention group (N=5/8%)
[2] Total number of events is less than 300 (a threshold rule-of-thumb)
[3] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
[4] Paper omits data
[5] Total population size is less than 400 (a threshold rule-of-thumb)

Infant emotional development: mother-infant relationship interventions versus treatment as usual

There was single study (N=63) evidence for a large harm associated with a mother–infant relationship intervention for women who had had a preterm delivery on preventing infant social withdrawal with infants in the intervention group showing worse scores than infants whose mothers had received treatment as usual ($p<0.00001$). However, confidence in this effect estimate was very low due to the very small sample size and the high risk of selective reporting bias. In addition, clinical and statistical significance of this effect estimate were not maintained at short-term follow-up ($p=0.59$) (Table 80).

Another study (N=84) found no evidence for clinically or statistically significant effects of a mother–infant relationship intervention for mothers who had had a preterm delivery on preventing problems with infant social-communication development ($p=0.88$) (Table 80).

Table 80: Summary of findings table for effects of mother–infant relationship interventions compared with treatment as usual on preventing emotional development problems in infants where mothers have identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Infant emotional development: Mother-infant relationship interventions versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Infant social-communication development mean scores post-treatment – available case analysis (at-risk populations) Pictoral Infant Communication Scales (PICS) Follow-up: mean 53 weeks		The mean infant social-communication development mean scores post-treatment – available case analysis (at-risk populations) in the intervention groups was 0.03 standard deviations higher (0.4 lower to 0.47 higher)		82 (1 study)	⊕⊖⊖⊖ very low[1,2]	SMD 0.03 (-0.4 to 0.47)
Infant social withdrawal mean scores post-treatment – available case analysis (at-risk populations) STSI: Approach Follow-up: mean 15 weeks		The mean infant social withdrawal mean scores post-treatment – available case analysis (at-risk populations) in the intervention groups was 1.52 standard deviations higher (0.95 to 2.08 higher)		63 (1 study)	⊕⊖⊖⊖ very low[2,3]	SMD 1.52 (0.95 to 2.08)
Infant social withdrawal mean scores Short follow-up (9-16 weeks post-intervention) – available case analysis (at-risk populations) STSI: Approach Follow-up: mean 28 weeks		The mean infant social withdrawal mean scores short follow-up (9-16 weeks post-intervention) – available case analysis (at-risk populations) in the intervention groups was 0.14 standard deviations higher (0.36 lower to 0.63 higher)		63 (1 study)	⊕⊖⊖⊖ very low[1,2,3,4]	SMD 0.14 (-0.36 to 0.63)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the

comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] High risk of selection bias due to statistically significant baseline difference with the intervention group having more mothers with earlier preterm birth and non-Norwegian origin
[2] Total population size is less than 400 (a threshold rule-of-thumb)
[3] Paper omits data
[4] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

7.4.19 Clinical evidence for effects on prevention of neglect or abuse of the infant where mothers have identified risk factors for mental health problems (by intervention)

Summary of findings can be found in the tables presented in this section. The full GRADE evidence profiles and associated forest plots can be found in Appendix 22 and Appendix 19, respectively.

Prevention of neglect or abuse of the infant: home visits versus treatment as usual

A single study (N=131) found evidence for large effects of home visits for women with psychosocial risk factors and (family) history of mental health problems on increasing the incidence of children being removed from the home (p=0.15) but reducing infant mortality (p=0.47). However, neither effect estimate was statistically significant due to very serious imprecision. The same study found no evidence for a clinically or statistically significant effect of home visits on preventing child protection issues (p=0.60). Another study (N=79) reported effects of home visits for adolescent mothers with psychosocial risk factors on preventing neglect or abuse of the infant, however, it was not possible to calculate an effect size due to zero cell counts (Table 81).

Table 81: Summary of findings table for effects of home visits compared with treatment as usual for prevention of neglect or abuse of the infant where mothers have identified risk factors for mental health problems

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	 Corresponding risk Prevention of neglect or abuse of the infant: Home visits versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Child protection issues Post-treatment – ITT analysis (at-risk populations) Follow-up: mean 78 weeks	Study population		RR 1.24 (0.56 to 2.73)	131 (1 study)	⊕⊕⊖⊖ low[1,2]	
	143 per 1000	177 per 1000 (80 to 390)				
	Moderate					
	143 per 1000	177 per 1000 (80 to 390)				
Child removed from home Post-treatment – ITT analysis (at-risk populations) Follow-up: mean 78 weeks	Study population		RR 8.35 (0.46 to 152)	131 (1 study)	⊕⊕⊖⊖ low[1,2]	
	0 per 1000	0 per 1000 (0 to 0)				
	Moderate					
	0 per 1000	0 per 1000 (0 to 0)				
Infant mortality Post-treatment – ITT analysis (at-risk populations) Follow-up: mean 78 weeks	Study population		RR 0.31 (0.01 to 7.45)	131 (1 study)	⊕⊕⊖⊖ low[1,2]	
	16 per 1000	5 per 1000 (0 to 118)				
	Moderate					
	16 per 1000	5 per 1000 (0 to 119)				
Infant abuse or neglect Post-treatment – available case analysis (at-risk populations)	See comment	See comment	Not estimable	79 (1 study)	See comment	

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total number of events is less than 300 (a threshold rule-of-thumb)
[2] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

7.4.20 Protocols for women following stillbirth

Depression for women who saw and/or held versus did not see and/or hold their stillborn infant

There was single study (N=65) data for large harms associated with seeing the stillborn infant for depression symptomatology during a subsequent pregnancy (p=0.08) and at one-year post-subsequent pregnancy follow-up (p=0.52). However, these effect estimates were imprecise due to low event rates and the 95% CI included no effect, appreciable benefit and appreciable harm. Another study with a much larger sample size (N=295) found no evidence for clinically or statistically significant harms associated with seeing (or not seeing) the stillborn infant on depression symptomatology 3 years post-stillbirth (p=0.59). Effects on depression mean symptoms were also not clinically or statistically significant (p=0.12-0.22) (Table 82).

The pattern of results was similar for depression outcomes associated with holding the stillborn infant, with single study (N=65) data for increased depression symptomatology during a subsequent pregnancy (p=0.03) or one-year post-subsequent pregnancy follow-up (p=0.16) associated with holding their stillborn infant. However, as before there are problems with imprecision of effect estimates and a larger study (N=295) found no evidence for increased risk of depression symptomatology 3-years post-stillbirth associated with holding (or not holding) their stillborn infant (p=0.99) (Table 82).

There was single study evidence for large benefits on depression symptomatology 3-years post-stillbirth of spending as much time with their stillborn infant as the woman wished (N=245; p <0.00001) but no evidence for clinically or statistically significant benefits or harms for depression symptomatology of keeping a photo of their stillborn infant (p=0.88), keeping a token of remembrance (p=0.51), or taking a drug to stop milk production following stillbirth (p=0.96) (Table 82).

Table 82: Summary of findings table for effects of seeing and/or holding and keeping mementoes compared with not seeing and/or holding the stillborn infant or keeping mementoes on depression outcomes

	Depression symptomatology	**Depression mean symptoms**
Study ID	(1)-(2) HUGHES2002/ TURTON2009 (3) RADESTAD2009A/ SURKAN2008	(1)-(2) HUGHES2002/ TURTON2009
Subgroup	(1)-(2) Pregnant at participation (3) Unclear pregnancy status at participation	(1)-(2) Pregnant at participation

Gestational age at loss (based on inclusion criteria)	(1)-(2) >18 weeks (3) >28 weeks	(1)-(2) >18 weeks
Time point	(1) During subsequent pregnancy (2) 1 year post-subsequent pregnancy follow-up (3) 3 years post-stillbirth	(1) During subsequent pregnancy (2) 1 year post-subsequent pregnancy follow-up
Outcome measure	(1) EPDS >14 (2) BDI >10 (3) CES-D >90th percentile	(1) EPDS (2) BDI
Number of studies (number of participants)	(1)-(2) K=1; N=65 (3) K=1; N=295	(1)-(2) K=1; N=65
Effect estimate for seeing the stillborn infant	(1) RR 5.67 [0.81, 39.55] (2) RR 1.59 [0.38, 6.65] (3) RR 0.76 [0.28, 2.05]	(1) SMD 0.44 [-0.12, 1.00] (2) SMD 0.35 [-0.21, 0.90]
Effect estimate for holding the stillborn infant	(1) RR 2.96 [1.08, 8.13] (2) RR 2.43 [0.71, 8.36] (3) RR 1.01 [0.48, 2.13]	(1) SMD 0.48 [-0.02, 0.97] (2) SMD 0.42 [-0.07, 0.91]
Effect estimate for spending as much time with stillborn infant as wished	(1)-(2) NR (3) RR 0.18 [0.09, 0.38]	(1)-(2) NR
Effect estimate for keeping a photo	(1)-(2) NR (3) RR 0.90 [0.23, 3.48]	(1)-(2) NR
Effect estimate for keeping a token of remembrance	(1)-(2) NR (3) RR 0.77 [0.36, 1.66]	(1)-(2) NR
Effect estimate for taking a drug to stop milk production following stillbirth	(1)-(2) NR (3) RR 0.95 [0.14, 6.23]	(1)-(2) NR
Note.		

Anxiety for women who saw and/or held versus did not see and/or hold their stillborn infant

There was single-study (N=65) evidence for clinically but not statistically significant harms of seeing or holding their stillborn infant on anxiety symptomatology during a subsequent pregnancy (p=0.19-0.21) or one-year post-subsequent pregnancy follow-up (p=0.08-0.64). This study also found a clinically and statistically significant moderate harm of seeing or holding the stillborn infant on mean anxiety symptoms during a subsequent pregnancy (p=0.03-0.05) though not at 1 year following the subsequent pregnancy (p=0.09-0.54). However, a larger single study (N=293) found no evidence for clinically or statistically significant harms (or benefits) of holding the stillborn infant on anxiety symptomatology 3-years post-stillbirth (p=0.73) (Table 83).

Table 83: Summary of findings table for effects of seeing and/or holding compared with not seeing and/or holding the stillborn infant on anxiety outcomes

	Anxiety symptomatology	**Anxiety mean symptoms**
Study ID	(1)-(2) HUGHES2002/ TURTON2009 (3) RADESTAD2009A/ SURKAN2008	(1)-(2) HUGHES2002/ TURTON2009
Subgroup	(1)-(2) Pregnant at participation (3) Unclear pregnancy status at participation	(1)-(2) Pregnant at participation
Gestational age at loss (based on inclusion criteria)	(1)-(2) >18 weeks (3) >28 weeks	(1)-(2) >18 weeks
Time point	(1) During subsequent pregnancy (2) 1 year post-subsequent pregnancy follow-up (3) 3 years post-stillbirth	(1) During subsequent pregnancy (2) 1 year post-subsequent pregnancy follow-up
Outcome measure	(1)-(2) STAI-S >44 (3) STAI-S >90th percentile	(1)-(2) STAI-S
Number of studies (number of participants)	(1)-(2) K=1; N=65 (3) K=1; N=293	(1)-(2) K=1; N=65
Effect estimate for seeing the stillborn infant	(1) RR 2.01 [0.67, 6.00] (2) RR 1.42 [0.33, 6.02] (3) NR	(1) SMD 0.64 [0.08, 1.21] (2) SMD 0.17 [-0.38, 0.73]
Effect estimate for holding the stillborn infant	(1) RR 1.69 [0.78, 3.69] (2) RR 3.65 [0.84, 15.88] (3) RR 0.89 [0.46, 1.71]	(1) SMD 0.50 [0.01, 1.00] (2) SMD 0.43 [-0.06, 0.92]
Note.		

PTSD for women who saw and/or held versus did not see and/or hold their stillborn infant

There was single study (N=65) evidence for a large and harmful effect of seeing the stillborn infant on PTSD symptomatology during a subsequent pregnancy (p=0.15). However, this effect estimate is imprecise due to the optimal information size (events=300) not being met and the 95% CI includes no effect, appreciable benefit and appreciable harm. This study also found a large harmful effect of seeing the stillborn infant on mean PTSD symptoms one-year post-subsequent pregnancy follow-up (p=0.003) but not during the subsequent pregnancy (p=0.16). This study also found large harms associated with holding the stillborn infant on PTSD symptomatology during a subsequent pregnancy (p=0.07), and large to moderate harms of holding the stillborn infant for mean PTSD symptoms during a subsequent pregnancy (p=0.02) and at 1-year (p=0.0002) and 7-year (p=0.009) post-subsequent pregnancy follow-ups . However, another study (N=98) found large benefits associated with holding the stillborn infant on PTSD symptomatology 5-18 years post-stillbirth (p=0.0009) (Table 84).

Table 84: Summary of findings table for effects of seeing and/or holding compared with not seeing and/or holding the stillborn infant on PTSD outcomes

	PTSD symptomatology	PTSD mean symptoms
Study ID	(1) HUGHES2002/ TURTON2009 (2) GRAVENSTEEN2013	(1)-(3) HUGHES2002/ TURTON2009
Subgroup	(1) Pregnant at participation (2) Not pregnant at participation	(1)-(3) Pregnant at participation
Gestational age at loss (based on inclusion criteria)	(1) >18 weeks (2) ≥23 weeks	(1)-(3) >18 weeks
Time point	(1) During subsequent pregnancy (2) 5-18 years post-stillbirth	(1) During subsequent pregnancy (2) 1 year post-subsequent pregnancy follow-up (3) 7 years post-subsequent pregnancy follow-up
Outcome measure	(1) PTSD-1 (DSM-III-R criteria) (2) IES >20	(1)-(2) PTSD-1
Number of studies (number of participants)	(1) K=1; N=65 (2) K=1; N=98	(1)-(2) K=1; N=65 (3) K=1; N=52
Effect estimate for seeing the stillborn infant	(1) RR 4.25 [0.60, 30.28] (2) NR	(1) SMD 0.40 [-0.16, 0.96] (2) SMD 0.88 [0.31, 1.46] (3) NR
Effect estimate for holding the stillborn infant	(1) RR 3.04 [0.92, 10.04] (2) RR 0.41 [0.24, 0.69]	(1) SMD 0.58 [0.09, 1.08] (2) SMD 1.00 [0.48, 1.52] (3) SMD 0.77 [0.19, 1.34]
Note.		

Summary of evidence for protocols for women following stillbirth

The evidence for benefits or harms associated with seeing and/or holding the stillborn infant was contradictory with evidence from HUGHES2002/TURTON2009 suggestive of harms associated with these protocols following stillbirth and evidence from RADESTAD2009A/SURKAN2008 and GRAVENSTEEN2013 suggestive of benefits associated with spending as much time with the stillborn infant as women wished or holding the stillborn infant. In addition, data could not be extracted for CACCIATORE2008 but narrative review of this study is consistent with the equivocal findings. Unfortunately, there is insufficient data to allow for sub-analyses. However, potential reasons for these differences could be differences in gestational age at the time of stillbirth. None of the papers report the mean gestational age at stillbirth, however, differences in the inclusion criteria are potentially consistent with more negative effects associated with these protocols for stillbirths occurring at earlier gestational ages (for instance, the inclusion criteria for HUGHES2002/TURTON2009 is >18 weeks compared to the inclusion criteria for RADESTAD2009A/SURKAN2008 which is >28 weeks). Another potential confounding factor and possible explanation for the mixed results is pregnancy

status at the time of participation in the studies and more negative effects associated with seeing and/or holding the stillborn infant observed during a subsequent pregnancy (as in HUGHES2002/TURTON2009) as compared to women who were not pregnant at the time of the study (as in GRAVENSTEEN2013). Narrative review of CACCIATORE2008 supports the hypothesis that pregnancy status may account for some of the between-study differences as that study found that seeing and/or holding their stillborn infant was associated with lower levels of depression for women who were non-pregnant when completing the questionnaire, while for women who were pregnant subsequent to a stillbirth seeing and/or holding was associated with a tendency towards depression.

7.4.21 Studies considered (prevention: no identified risk factors)

Seven RCTs reported across ten papers met the eligibility criteria for this review: HOWELL2014 (Howell et al., 2014), KALINAUSKIENE2009 (Kalinauskiene et al., 2009), LAVENDER1998 (Lavender & Walkinshaw, 1998), MORRELL2000 (Morrell et al., 2000), MORRELL2009A/2009B/2011/BRUGHA2011 (Morrell et al., 2009a; Morrell et al., 2009b; Morrell et al., 2011; Brugha et al., 2011), PEREZBLASCO2013 (Perez-Blasco et al., 2013), TSENG2010 (Tseng et al., 2010). All of these studies were published in peer-reviewed journals between 1998 and 2013. In addition, 28 studies were excluded from the review. The most common reasons for exclusion were that data could not be extracted, there were no mental health outcomes reported, the group assignment was non-randomised, or the intervention was outside the scope (for instance, organisation of care trials). Further information about both included and excluded studies can be found in Appendix 18.

Of the seven included RCTs, there was one study (N=2,324) involving a comparison of a structured psychological intervention (CBT) and treatment as usual (Table 85).

There was one study (N=2,297) that compared listening visits with treatment as usual (Table 86).

There were two studies (N=1,978) that involved a comparison between psychologically (CBT/IPT)-informed psychoeducation and enhanced treatment as usual, one study (N=623) involved a comparison of home visits and treatment as usual, and one study (N=120) compared post-delivery discussion and treatment as usual (Table 87).

One study (N=54) compared a mother–infant relationship intervention and enhanced treatment as usual (Table 88). Although the participants in this study did not meet criteria for the pre-specified risk factors, the mothers were classified as 'insensitive' at baseline (defined as score<5 [midpoint] on Ainsworth rating scale for sensitivity).

Finally, there was one study (N=92) that involved a comparison between music therapy and treatment as usual and one study (N=26) compared mindfulness training with treatment as usual (Table 89).

Table 85: Study information table for trials included in the prevention (no risk factors identified) meta-analysis of structured psychological interventions (CBT or IPT) versus any alternative management strategy

	Structured psychological interventions (CBT or IPT) versus TAU
Total number of trials (number of participants)	1 (2,324)
Study ID	MORRELL2009A/2009B/2011/BRUGHA2011[1]
Country	UK
Mean age of participants (years)	31.5
Timing of intervention	Postnatal
Mode of delivery	Face-to-face
Format	Individual
Intensity (number of sessions)	Moderate (8 sessions)
Length of intervention (weeks)	8
Time points	First measurement
Setting	Home
Intervention	CBT
Comparison	TAU
Note. Abbreviations: TAU=Treatment as usual [1]Three-armed trial that includes both prevention (whole sample) and treatment ('depressed' subgroup) data: CBT; Listening visits; TAU. Listening visits versus TAU comparison extracted below. Demographic data is based on all three arms.	

Table 86: Study information table for trials included in the prevention (no risk factors identified) meta-analysis of counselling versus any alternative management strategy

	Listening visits versus TAU
Total number of trials (number of participants)	1 (2,297)
Study ID	MORRELL2009A/2009B/2011/BRUGHA2011[1]
Country	UK
Mean age of participants (years)	31.5
Timing of intervention	Postnatal
Mode of delivery	Face-to-face
Format	Individual
Intensity (number of sessions)	Moderate (8 sessions)
Length of intervention (weeks)	8
Time points	First measurement
Setting	Home

Intervention	Listening visits ('person centred approach')
Comparison	TAU

Note. Abbreviations: TAU=Treatment as usual
[1]Three-armed trial that includes both prevention (whole sample) and treatment ('depressed' subgroup) data: CBT; Listening visits; TAU. CBT versus TAU comparison extracted above. Demographic data is based on all three arms.

Table 87: Study information table for trials included in the prevention (no risk factors identified) meta-analysis of education and support versus any alternative management strategy

	Psychologically (CBT/IPT)-informed psychoeducation versus Enhanced TAU	**Home visits versus TAU**	**Post-delivery discussion versus TAU**
Total number of trials (number of participants)	2 (1,978)	1 (623)	1 (120)
Study ID	(1) HOWELL2014 (2) KOZINSZKY2012[1]	MORRELL2000	LAVENDER1998
Country	(1) US (2) Hungary	UK	UK
Mean age of participants (years)	(1) 32.5 (2) 27.3	27.8	24.2
Timing of intervention	(1) Postnatal (2) Antenatal	Postnatal	Postnatal
Mode of delivery	(1) Booklet (with face-to-face support) and telephone (2) Face-to-face	Face-to-face	Face-to-face
Format	(1) Individual (2) Group	Individual	Individual
Intensity (number of sessions)	(1) Low (2 sessions) (2) Low (4 sessions)	Low (6 sessions)	Low (single session)
Length of intervention (weeks)	(1) 2 (2) 4	4	Single session
Time points	(1) Post-treatment; Short follow-up; Intermediate follow-up (2) First measurement	Post-treatment; Intermediate follow-up	Post-treatment
Setting	(1) Hospital and telephone (2) NR	Home	Hospital
Intervention	(1) Behavioural educational intervention (2) Psychologically-informed psychoeducation group sessions	Home visits	Debriefing
Comparison	(1) Enhanced TAU (non-mental health-focused	TAU	TAU

	education and support [booklet and telephone call]) (2) Enhanced TAU (non-mental health-focused education and support [group])		
Note. Abbreviations: NR=Not reported; TAU=Treatment as usual [1]Paper also reports data for a 'depressed' subgroup which is extracted in the treatment section below			

Table 88: Study information table for trials included in the prevention (no risk factors identified) meta-analysis of mother–infant relationship interventions versus any alternative management strategy

	Mother-infant relationship interventions versus Enhanced TAU
Total number of trials (number of participants)	1 (54)
Study ID	KALINAUSKIENE2009
Country	Lithuania
Mean age of participants (years)	26.4
Timing of intervention	Postnatal
Mode of delivery	Face-to-face
Format	Individual
Intensity (number of sessions)	Low (5 sessions)
Length of intervention (weeks)	22
Time points	Post-treatment
Setting	Home
Intervention	Video-feedback intervention to promote positive parenting (VIPP)
Comparison	Enhanced TAU (monitoring)
Note. Abbreviations: TAU=Treatment as usual	

Table 89: Study information table for trials included in the prevention (no risk factors identified) meta-analysis of other psychosocial interventions versus any alternative management strategy

	Music therapy versus TAU	**Mindfulness training versus TAU**
Total number of trials (number of participants)	1 (92)	1 (26)
Study ID	TSENG2010	PEREZBLASCO2013
Country	Taiwan	Spain
Mean age of participants (years)	30.6	34.3
Timing of intervention	Postnatal	Postnatal
Mode of delivery	CD	Face-to-face

Format	Individual	Group
Intensity (number of sessions)	Low (0 contact with professionals [14 CD sessions])	Moderate (8 sessions)
Length of intervention (weeks)	2	8
Time points	Post-treatment	Post-treatment
Setting	Home	Clinic (primary)
Intervention	Music therapy	Mindfulness-based intervention
Comparison	TAU	Waitlist
Note. Abbreviations: TAU=Treatment as usual		

7.4.22 Clinical evidence for preventative effects on depression outcomes for women with no identified risk factors (by intervention)

Summary of findings can be found in the tables presented in this section. The full GRADE evidence profiles and associated forest plots can be found in Appendix 22 and Appendix 19, respectively.

Depression: structured psychological interventions (CBT or IPT) versus treatment as usual

There was single study (N=1,762) available case analysis evidence for a moderate effect of CBT relative to treatment as usual for preventing depression symptomatology in women in the postnatal period with no identified risk factors (p=0.004). However, the ITT analysis of the same outcome measure showed no evidence of statistically or clinically significant preventative effects (p=0.97). There was also no evidence for a clinically significant effect (although it was statistically significant [p <0.00001]) on mean depression symptoms (Table 90).

Table 90: Summary of findings table for effects of structured psychological interventions (CBT or IPT) compared with treatment as usual on preventing depression outcomes in women with no identified risk factors

Outcomes	**Illustrative comparative risks* (95% CI)** **Assumed Corresponding risk risk** **Control** **Depression: Structured psychological interventions (CBT or IPT) versus TAU**	**Relative effect (95% CI)**	**No. of Participants (studies)**	**Quality of the evidence (GRADE)**	**Comments**
	Study population				

Depression symptomatology Post-treatment – ITT analysis (no-risk populations) EPDS ≥12 Follow-up: mean 26 weeks	348 per 1000	348 per 1000 (313 to 390)	RR 1 (0.9 to 1.12)	2,324 (1 study)	⊕⊕⊕⊖ moderate[1]	
	Moderate					
	348 per 1000	348 per 1000 (313 to 390)				
Depression symptomatology Post-treatment – available case analysis (no-risk populations) EPDS ≥12 Follow-up: mean 26 weeks	Study population		RR 0.7 (0.56 to 0.89)	1,762 (1 study)	⊕⊕⊖⊖ low[1,2]	
	164 per 1000	115 per 1000 (92 to 146)				
	Moderate					
	164 per 1000	115 per 1000 (92 to 146)				
Depression mean scores post-treatment – available case analysis (no-risk populations) EPDS Follow-up: mean 26 weeks		The mean depression mean scores post-treatment – available case analysis (no-risk populations) in the intervention groups was 0.22 standard deviations lower (0.31 to 0.13 lower)		1,762 (1 study)	⊕⊕⊕⊖ moderate[1]	SMD -0.22 (-0.31 to -0.13)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Paper omits data
[2] Total number of events is less than 300 (a threshold rule-of-thumb)

Depression: listening visits versus treatment as usual

Using an available case analysis approach, there was single study (N=1,811) evidence for a moderate preventative effect of listening visits on depression symptomatology for women in the postnatal period with no identified risk factors (p=0.007). However, the ITT analysis for depression symptomatology revealed no clinically significant difference between listening visits and treatment as usual, although the difference was statistically significant (p=0.01). For depression mean scores there was also a statistically significant (p <0.0001) but not an appreciable benefit of listening visits (Table 91).

Table 91: Summary of findings table for effects of listening visits compared with treatment as usual on preventing depression outcomes in women with no identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Depression: Listening visits versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Depression symptomatology Post-treatment – ITT analysis (no-risk populations) EPDS ≥12 Follow-up: mean 26 weeks	Study population 348 per 1000	299 per 1000 (265 to 334) Moderate 348 per 1000	RR 0.86 (0.76 to 0.96) 299 per 1000 (264 to 334)	2,297 (1 study)	⊕⊕⊕⊖ moderate[1]	
Depression symptomatology Post-treatment – available case analysis (no-risk populations) EPDS ≥12 Follow-up: mean 26 weeks	Study population 164 per 1000	120 per 1000 (95 to 151) Moderate 164 per 1000	RR 0.73 (0.58 to 0.92) 120 per 1000 (95 to 151)	1,811 (1 study)	⊕⊕⊖⊖ low[1,2]	
Depression mean scores post-treatment – available case analysis (no-risk populations) EPDS Follow-up: mean 26 weeks		The mean depression mean scores post-treatment – available case analysis (no-risk populations) in the intervention groups was 0.2 standard deviations lower (0.3 to 0.11 lower)		1,811 (1 study)	⊕⊕⊕⊖ moderate[1]	SMD -0.2 (-0.3 to -0.11)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Paper omits data
[2] Total number of events is less than 300 (a threshold rule-of-thumb)

Depression: Psychologically (CBT/IPT)-informed psychoeducation versus enhanced treatment as usual

There was no evidence for statistically or clinically significant benefits of psychoeducation for preventing depression in the postnatal period for women with no identified risk factors (p=0.51-0.99; Table 92).

Table 92: Summary of findings table for effects of psychologically (CBT/IPT)-informed psychoeducation compared with enhanced treatment as usual on preventing depression outcomes in women with no identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk / Corresponding risk Control / Depression: Psychologically (CBT/IPT)-informed psychoeducation versus Enhanced TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Depression symptomatology Post-treatment – ITT analysis (no-risk populations) EPDS ≥10 or Leverton Questionnaire (Elliott et al., 2000) ≥12 Follow-up: 4-17 weeks	Study population 100 per 1000 / 100 per 1000 (77 to 131) Moderate 108 per 1000 / 108 per 1000 (83 to 141)	RR 1 (0.77 to 1.31)	1,978 (2 studies)	⊕⊕⊖⊖ low1,2	
Depression symptomatology Post-treatment – available case analysis (no-risk populations) EPDS ≥10 Follow-up: mean 4 weeks	Study population 56 per 1000 / 60 per 1000 (30 to 122) Moderate 56 per 1000 / 60 per 1000 (30 to 123)	RR 1.08 (0.53 to 2.19)	500 (1 study)	⊕⊕⊖⊖ low1,2	
Depression symptomatology Short Follow-up (9-16 weeks post-intervention) – ITT analysis (no-risk populations) EPDS ≥10 Follow-up: mean 12 weeks	Study population 196 per 1000 / 175 per 1000 (122 to 247) Moderate 196 per 1000 / 174 per 1000 (122 to 247)	RR 0.89 (0.62 to 1.26)	540 (1 study)	⊕⊕⊖⊖ low1,2	
Depression symptomatology Short Follow-up (9-16 weeks	Study population 65 per 1000 / 51 per 1000 (25 to 107)	RR 0.79 (0.38 to 1.65)	467 (1 study)	⊕⊕⊖⊖ low1,2	

post-intervention) – available case analysis (no-risk populations) EPDS ≥10 Follow-up: mean 12 weeks	Moderate 65 per 1000 — 51 per 1000 (25 to 107)			
Depression symptomatology Intermediate follow-up (17-24 weeks post-intervention) – ITT analysis (no-risk populations) EPDS ≥10 Follow-up: mean 25 weeks	Study population 159 per 1000 — 178 per 1000 (123 to 258) Moderate 159 per 1000 — 178 per 1000 (122 to 258)	RR 1.12 (0.77 to 1.62)	540 (1 study)	⊕⊕⊖⊖ low1,2
Depression symptomatology Intermediate follow-up (17-24 weeks post-intervention) – available case analysis (no-risk populations) EPDS ≥10 Follow-up: mean 25 weeks	Study population 46 per 1000 — 35 per 1000 (14 to 85) Moderate 46 per 1000 — 34 per 1000 (14 to 85)	RR 0.75 (0.31 to 1.84)	468 (1 study)	⊕⊕⊖⊖ low1,2

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total number of events is less than 300 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Depression: home visits versus treatment as usual

There was no evidence for statistically or clinically significant benefits of home visits relative to treatment as usual for reducing mean depression symptoms at 6 weeks (p=0.13) or 6 months (p=0.84) postnatally for women with no identified risk factors (Table 93).

Table 93: Summary of findings table for effects of home visits compared with treatment as usual on preventing depression outcomes in women with no identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Depression: Home visits versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Depression mean scores post-treatment – available case analysis (no-risk populations) EPDS Follow-up: mean 6 weeks		The mean depression mean scores post-treatment – available case analysis (no-risk populations) in the intervention groups was 0.13 standard deviations higher (0.04 lower to 0.3 higher)		542 (1 study)	⊕⊕⊕⊖ moderate[1]	SMD 0.13 (-0.04 to 0.3)
Depression mean scores Intermediate follow-up (17-24 weeks post-intervention) – available case analysis (no-risk populations) EPDS Follow-up: mean 26 weeks		The mean depression mean scores intermediate follow-up (17-24 weeks post-intervention) – available case analysis (no-risk populations) in the intervention groups was 0.02 standard deviations lower (0.2 lower to 0.16 higher)		481 (1 study)	⊕⊕⊕⊖ moderate[1]	SMD -0.02 (-0.2 to 0.16)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Risk of bias due to statistically significant group differences at baseline

Depression: post-delivery discussion versus treatment as usual

There was single study (N=114) evidence for a large effect of post-delivery discussion relative to treatment as usual for preventing depression symptomatology in the postnatal period for women with no identified risk factors ($p < 0.0001$). However, the confidence in this effect estimate is low due to very serious imprecision as the optimal information size (events=300) is not met (Table 94).

Table 94: Summary of findings table for effects of post-delivery discussion compared with treatment as usual on preventing depression outcomes in women with no identified risk factors

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Depression: post-delivery discussion versus TAU				
Depression symptomatology Post-treatment – available case analysis (no-risk populations) HADS – Depression ≥11 Follow-up: mean 3 weeks	Study population		RR 0.16 (0.07 to 0.37)	114 (1 study)	⊕⊕⊖⊖ low[1]	
	554 per 1000	89 per 1000 (39 to 205)				
	Moderate					
	554 per 1000	89 per 1000 (39 to 205)				

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total number of events is less than 300 (a threshold rule-of-thumb)

Depression: mother-infant relationship interventions versus enhanced treatment as usual

There was no evidence for statistically or clinically significant benefits of mother–infant relationship interventions relative to monitoring for reducing mean depression symptoms in the postnatal period for women with no identified risk factors (p=0.32; Table 95).

Table 95: Summary of findings table for effects of mother–infant relationship interventions compared with enhanced treatment as usual on preventing depression outcomes in women with no identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Depression: Mother-infant relationship interventions versus Enhanced TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Depression mean scores post-treatment – ITT analysis (no-risk populations) BDI Follow-up: mean 26 weeks		The mean depression mean scores post-treatment – ITT analysis (no-risk populations) in the intervention groups was 0.27 standard deviations lower (0.81 lower to 0.26 higher)		54 (1 study)	⊕⊕⊖⊖ low[1,2]	SMD -0.27 (-0.81 to 0.26)
Depression mean scores post-treatment – available case analysis (no-risk populations) BDI Follow-up: mean 26 weeks		The mean depression mean scores post-treatment – available case analysis (no-risk populations) in the intervention groups was 0.27 standard deviations lower (0.81 lower to 0.26 higher)		54 (1 study)	⊕⊕⊖⊖ low[1,2]	SMD -0.27 (-0.81 to 0.26)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total population size is less than 400 (a threshold rule-of-thumb)
[2] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Depression: mindfulness training versus treatment as usual

There was no evidence for statistically or clinically significant benefits of mindfulness training relative to treatment as usual for reducing depression mean symptoms in the postnatal period for women with no identified risk factors (p=0.42; Table 96).

Table 96: Summary of findings table for effects of mindfulness training compared with treatment as usual on preventing depression outcomes in women with no identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Depression: Mindfulness training versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Depression mean scores post-treatment – available case analysis (no-risk populations) Depression, Anxiety, and Stress Scale (DASS-21): Depression Follow-up: mean 11 weeks		The mean depression mean scores post-treatment – available case analysis (no-risk populations) in the intervention groups was 0.36 standard deviations lower (1.25 lower to 0.53 higher)		21 (1 study)	⊕⊕⊖⊖ low1,2	SMD -0.36 (-1.25 to 0.53)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total population size is less than 400 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

7.4.23 Clinical evidence for preventative effects on anxiety outcomes for women with no identified risk factors (by intervention)

Summary of findings can be found in the tables presented in this section. The full GRADE evidence profiles and associated forest plots can be found in Appendix 22 and Appendix 19, respectively.

Anxiety: structured psychological interventions (CBT or IPT) versus treatment as usual

There was no evidence for clinically significant benefits of CBT relative to treatment as usual for reducing anxiety symptoms (state and trait) in the postnatal period for women with no identified risk factors, although the effects were statistically significant (p=0.007-0.01) they were too small to be considered clinically meaningful (Table 97).

Table 97: Summary of findings table for effects of structured psychological interventions (CBT or IPT) compared with treatment as usual on preventing anxiety outcomes in women with no identified risk factors

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Anxiety: Structured psychological interventions (CBT or IPT) versus TAU				
Anxiety mean scores post-treatment – available case analysis (no-risk populations) STAI-S Follow-up: mean 26 weeks		The mean anxiety mean scores post-treatment – available case analysis (no-risk populations) in the intervention groups was 0.13 standard deviations lower (0.23 to 0.04 lower)		1,653 (1 study)	⊕⊕⊕⊖ moderate[1]	SMD -0.13 (-0.23 to -0.04)
Trait anxiety mean scores post-treatment – available case analysis (no-risk populations) STAI-T Follow-up: mean 26 weeks		The mean trait anxiety mean scores post-treatment – available case analysis (no-risk populations) in the intervention groups was 0.12 standard deviations lower (0.22 to 0.02 lower)		1,618 (1 study)	⊕⊕⊕⊖ moderate[1]	SMD -0.12 (-0.22 to -0.02)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Paper omits data

Anxiety: listening visits versus treatment as usual

Although statistically significant benefits of listening visits for reducing postnatal state and trait anxiety symptoms were observed (p=0.03-0.04), the effect sizes were too small to be considered as showing an appreciable clinical benefit (Table 98).

Table 98: Summary of findings table for effects of listening visits compared with treatment as usual on preventing anxiety outcomes in women with no identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	 Corresponding risk Anxiety: Listening visits versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Anxiety mean scores post-treatment – available case analysis (no-risk populations) STAI-S Follow-up: mean 26 weeks		The mean anxiety mean scores post-treatment – available case analysis (no-risk populations) in the intervention groups was 0.1 standard deviations lower (0.19 lower to 0 higher)		1,697 (1 study)	⊕⊕⊕⊖ moderate[1]	SMD -0.1 (-0.19 to 0)
Trait anxiety mean scores post-treatment – available case analysis (no-risk populations) STAI-T Follow-up: mean 26 weeks		The mean trait anxiety mean scores post-treatment – available case analysis (no-risk populations) in the intervention groups was 0.11 standard deviations lower (0.2 to 0.01 lower)		1,695 (1 study)	⊕⊕⊕⊖ moderate[1]	SMD -0.11 (-0.2 to -0.01)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.

Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Paper omits data

Anxiety: post-delivery discussion versus treatment as usual

There was single study (N=114) evidence for a large effect of a post-delivery discussion on preventing anxiety symptomatology in the postnatal period for women with no identified risk factors ($p < 0.0001$). However, the confidence in this effect estimate is low due to very serious imprecision conferred by a low event rate (Table 99).

Table 99: Summary of findings table for effects of post-delivery discussion compared with treatment as usual on preventing anxiety outcomes in women with no identified risk factors

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Anxiety: post-delivery discussion versus TAU				
Anxiety symptomatology Post-treatment – available case analysis (no-risk populations) HADS – Anxiety ≥11 Follow-up: mean 3 weeks	Study population		RR 0.14 (0.05 to 0.37)	114 (1 study)	⊕⊕⊖⊖ low[1]	
	500 per 1000	70 per 1000 (25 to 185)				
	Moderate					
	500 per 1000	70 per 1000 (25 to 185)				

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total number of events is less than 300 (a threshold rule-of-thumb)

Anxiety: music therapy versus treatment as usual

There was no evidence for statistically or clinically significant effects of music therapy for reducing anxiety symptoms in the postnatal period for women with no identified risk factors (p=0.07; Table 100).

Table 100: Summary of findings table for effects of music therapy compared with treatment as usual on preventing anxiety outcomes in women with no identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Anxiety: Music therapy versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Anxiety mean scores post-treatment – available case analysis (no-risk populations) STAI-S Follow-up: mean 2 weeks		The mean anxiety mean scores post-treatment – available case analysis (no-risk populations) in the intervention groups was 0.42 standard deviations higher (0.04 lower to 0.87 higher)		77 (1 study)	⊕⊖⊖⊖ very low1,2,3	SMD 0.42 (-0.04 to 0.87)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Risk of bias due to statistically significant group differences at baseline
2 Total population size is less than 400 (a threshold rule-of-thumb)
3 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Anxiety: mindfulness training versus treatment as usual

There was single study (N=21) evidence for a very large effect of mindfulness training on reducing anxiety symptoms in the postnatal period for women with no identified risk factors (p=0.01). However, confidence in this effect estimate was low due to very serious imprecision as a result of the very small sample size (Table 101).

Table 101: Summary of findings table for effects of mindfulness training compared with treatment as usual on preventing anxiety outcomes in women with no identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Anxiety: Mindfulness training versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Anxiety mean scores post-treatment – available case analysis (no-risk populations) DASS-21: Anxiety Follow-up: mean 11 weeks		The mean anxiety mean scores post-treatment – available case analysis (no-risk populations) in the intervention groups was 1.21 standard deviations lower (2.18 to 0.24 lower)		21 (1 study)	⊕⊕⊖⊖ low[1]	SMD -1.21 (-2.18 to -0.24)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total population size is less than 400 (a threshold rule-of-thumb)

7.4.24 Clinical evidence for preventative effects on poor general mental health outcomes for women with no identified risk factors (by intervention)

Summary of findings can be found in the tables presented in this section. The full GRADE evidence profiles and associated forest plots can be found in Appendix 22 and Appendix 19, respectively.

General mental health: structured psychological interventions (CBT or IPT) versus treatment as usual

There was single study (N=1,749) moderate quality evidence for a moderate benefit of CBT relative to treatment as usual, for women in the postnatal period with no identified risk factors, on lower risk of self-harm (Table 102). The same study

(N=1,700) found no clinically significant benefit (although the effect was statistically significant) of CBT on preventing poor general mental health mean scores (p=0.002).

Table 102: Summary of findings table for effects of structured psychological interventions (CBT or IPT) compared with treatment as usual on preventing poor general mental health outcomes in women with no identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk General mental health: Structured psychological interventions (CBT or IPT) versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
General mental health mean scores post-treatment – available case analysis (no-risk populations) Short Form Health Survey-12 (SF-12) Mental component summary Follow-up: mean 26 weeks		The mean general mental health mean scores post-treatment – available case analysis (no-risk populations) in the intervention groups was 0.15 standard deviations higher (0.06 to 0.25 higher)		1,700 (1 study)	⊕⊕⊕⊖ moderate[1]	SMD 0.15 (0.06 to 0.25)
Risk of self-harm mean scores post-treatment – available case analysis (no-risk populations) Clinical Outcomes in Routine Evaluation-Outcome Measure (CORE-OM): Risk of self-harm Follow-up: mean 26 weeks		The mean risk of self-harm mean scores post-treatment – available case analysis (no-risk populations) in the intervention groups was 0.66 standard deviations lower (0.75 to 0.56 lower)		1,749 (1 study)	⊕⊕⊕⊖ moderate[1]	SMD -0.66 (-0.75 to -0.56)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the

estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Paper omits data

General mental health: listening visits versus treatment as usual

There was single study (N=1,799) moderate quality evidence for a moderate benefit of listening visits relative to treatment as usual, for women in the postnatal period with no identified risk factors, on lower risk of self-harm (Table 103). The same study (N=1,764) found no clinically significant benefit (although the effect was statistically significant) of listening visits on preventing poor general mental health mean scores (p=0.001).

Table 103: Summary of findings table for effects of listening visits compared with treatment as usual on preventing poor general mental health outcomes in women with no identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk General mental health: Listening visits versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
General mental health mean scores post-treatment – available case analysis (no-risk populations) SF-12 Mental component summary Follow-up: mean 26 weeks		The mean general mental health mean scores post-treatment – available case analysis (no-risk populations) in the intervention groups was 0.15 standard deviations higher (0.06 to 0.25 higher)		1,764 (1 study)	⊕⊕⊕⊖ moderate[1]	SMD 0.15 (0.06 to 0.25)
Risk of self-harm mean scores post-treatment – available case analysis (no-risk populations) CORE-OM: Risk of self-harm Follow-up: mean 26 weeks		The mean risk of self-harm mean scores post-treatment – available case analysis (no-risk populations) in the intervention groups was 0.57 standard deviations higher (0.47 to 0.66 higher)		1,799 (1 study)	⊕⊕⊕⊖ moderate[1]	SMD 0.57 (0.47 to 0.66)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Paper omits data

General mental health: home visits versus treatment as usual

A single study (N=481-550) found no evidence for clinically or statistically significant effects of home visits for women in the postnatal period with no identified risk factors for preventing poor general mental health mean scores at post-treatment (p=0.64) or intermediate follow-up (p=0.45) (Table 104).

Table 104: Summary of findings table for effects of home visits compared with treatment as usual on preventing poor general mental health outcomes in women with no identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk General mental health: Home visits versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
General mental health mean scores post-treatment – available case analysis (no-risk populations) SF-36- Mental health Follow-up: mean 6 weeks		The mean general mental health mean scores post-treatment – available case analysis (no-risk populations) in the intervention groups was 0.04 standard deviations lower (0.21 lower to 0.13 higher)		550 (1 study)	⊕⊕⊕⊖ moderate[1]	SMD -0.04 (-0.21 to 0.13)
General mental health mean scores Intermediate follow-up (17-24 weeks post-intervention) – available case analysis (no-risk populations) SF-36- Mental health		The mean general mental health mean scores intermediate follow-up (17-24 weeks post-intervention) – available case analysis (no-risk populations) in the intervention groups was 0.07 standard deviations lower		481 (1 study)	⊕⊕⊕⊖ moderate[1]	SMD -0.07 (-0.25 to 0.11)

Follow-up: mean 26 weeks	(0.25 lower to 0.11 higher)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] High risk of selection bias due to statistically significant baseline group differences for incidence of twins, use of TENS during labour, and adults living with the mother

General mental health: mindfulness training versus treatment as usual

There was single study (N=21) evidence for a large effect of mindfulness training for women in the postnatal period with no identified risk factors on preventing psychological distress (p=0.02). However, confidence in this effect estimate is low due to very serious imprecision (very small sample size). The same study found no evidence for clinically or statistically significant effects of mindfulness training on life satisfaction (p=0.35) or happiness (p=0.60) (Table 105).

Table 105: Summary of findings table for effects of mindfulness training compared with treatment as usual on preventing poor general mental health outcomes in women with no identified risk factors

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	General mental health: Mindfulness training versus TAU				
Psychological distress mean scores post-treatment – available case analysis (no-risk populations) DASS-21: Psychological distress Follow-up: mean 11 weeks		The mean psychological distress mean scores post-treatment – available case analysis (no-risk populations) in the intervention groups was 1.15 standard deviations lower (2.11 to 0.19 lower)		21 (1 study)	⊕⊕⊖⊖ low[1]	SMD -1.15 (-2.11 to -0.19)

Life satisfaction mean scores post-treatment – available case analysis (no-risk populations) Satisfaction With Life Scale Follow-up: mean 11 weeks	The mean life satisfaction mean scores post-treatment – available case analysis (no-risk populations) in the intervention groups was 0.43 standard deviations higher (0.46 lower to 1.32 higher)	21 (1 study)	⨁⨁⊖⊖ low[1,2]	SMD 0.43 (-0.46 to 1.32)
Happiness mean scores post-treatment – available case analysis (no-risk populations) Subjective Happiness Scale Follow-up: mean 11 weeks	The mean happiness mean scores post-treatment – available case analysis (no-risk populations) in the intervention groups was 0.24 standard deviations higher (0.65 lower to 1.12 higher)	21 (1 study)	⨁⨁⊖⊖ low[1,2]	SMD 0.24 (-0.65 to 1.12)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total population size is less than 400 (a threshold rule-of-thumb)
[2] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

7.4.25 Clinical evidence for preventative effects on poor mental health outcomes for women with no identified risk factors (sub-analyses)

There was insufficient data to enable sub-analyses by treatment timing, format or intensity for the prevention (no risk factors identified) review.

7.4.26 Clinical evidence for preventative effects on mother–infant attachment problems for women with no identified risk factors (by intervention)

Summary of findings can be found in the tables presented in this section. The full GRADE evidence profiles and associated forest plots can be found in Appendix 22 and Appendix 19, respectively.

Mother-infant attachment: home visits versus treatment as usual

A single study (N=493-548) found no evidence for clinically or statistically significant effects of home visits for women in the postnatal period with no identified risk factors on preventing breastfeeding discontinuation before 6 weeks (p=0.50) or before 6 months (p=0.87) (Table 106).

Table 106: Summary of findings table for effects of home visits compared with treatment as usual on preventing mother–infant attachment problems for women with no identified risk factors

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Mother-infant attachment: Home visits versus TAU				
Discontinued breastfeeding by 6 weeks – available case analysis (no-risk populations) Follow-up: mean 6 weeks	Study population		RR 0.95 (0.82 to 1.1)	548 (1 study)	⊕⊕⊕⊖ moderate[1]	
	578 per 1000	549 per 1000 (474 to 636)				
	Moderate					
	578 per 1000	549 per 1000 (474 to 636)				
Discontinued breastfeeding by 26 weeks – available case analysis (no-risk populations) Follow-up: mean 26 weeks	Study population		RR 1.01 (0.92 to 1.1)	493 (1 study)	⊕⊕⊕⊖ moderate[1]	
	794 per 1000	802 per 1000 (730 to 873)				
	Moderate					
	794 per 1000	802 per 1000 (730 to 873)				

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] High risk of selection bias due to statistically significant baseline group differences for incidence of twins, use of TENS during labour, and adults living with the mother

Mother-infant attachment: mother-infant relationship intervention versus enhanced treatment as usual

There was single study (N=54) low quality evidence for a moderate effect of a mother–infant relationship intervention relative to enhanced treatment as usual (monitoring) for women in the postnatal period with no identified risk factors on preventing poor maternal sensitivity scores (p=0.007). However, this study found no clinically or statistically effects of a mother–infant relationship intervention on child attachment security (p=1.00) or maternal confidence/competence (p=0.28) (Table 107).

Table 107: Summary of findings table for effects of a mother–infant relationship intervention compared with enhanced treatment as usual on preventing mother–infant attachment problems for women with no identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	 Corresponding risk Mother-infant attachment: Mother-infant relationship interventions versus Enhanced TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Maternal sensitivity mean scores post-treatment – ITT analysis (no-risk populations) Ainsworth Strange Situation: Total Follow-up: mean 26 weeks		The mean maternal sensitivity mean scores post-treatment – ITT analysis (no-risk populations) in the intervention groups was 0.77 standard deviations higher (0.21 to 1.32 higher)		54 (1 study)	⊕⊕⊖⊖ low[1]	SMD 0.77 (0.21 to 1.32)
Maternal sensitivity mean scores post-treatment – available case analysis (no-risk populations) Ainsworth Strange Situation: Total Follow-up: mean 26 weeks		The mean maternal sensitivity mean scores post-treatment – available case analysis (no-risk populations) in the intervention groups was 0.77 standard deviations higher (0.21 to 1.32 higher)		54 (1 study)	⊕⊕⊖⊖ low[1]	SMD 0.77 (0.21 to 1.32)
Child attachment security mean scores post-treatment – ITT analysis (no-risk populations)		The mean child attachment security mean scores post-treatment – ITT analysis (no-risk		54 (1 study)	⊕⊕⊖⊖ low[1,2]	SMD 0 (-0.53 to 0.53)

Waters' Attachment Q-set Follow-up: mean 26 weeks	populations) in the intervention groups was 0 standard deviations higher (0.53 lower to 0.53 higher)			
Child attachment security mean scores post-treatment – available case analysis (no-risk populations) Waters' Attachment Q-set Follow-up: mean 26 weeks	The mean child attachment security mean scores post-treatment – available case analysis (no-risk populations) in the intervention groups was 0 standard deviations higher (0.53 lower to 0.53 higher)	54 (1 study)	⊕⊕⊖⊖ low[1,2]	SMD 0 (-0.53 to 0.53)
Maternal confidence/competence mean scores post-treatment – ITT analysis (no-risk populations) Parental Efficacy Questionnaire Follow-up: mean 26 weeks	The mean maternal confidence/competence mean scores post-treatment – ITT analysis (no-risk populations) in the intervention groups was 0.3 standard deviations higher (0.24 lower to 0.84 higher)	54 (1 study)	⊕⊕⊖⊖ low[1,2]	SMD 0.3 (-0.24 to 0.84)
Maternal confidence/competence mean scores post-treatment – available case analysis (no-risk populations) Parental Efficacy Questionnaire Follow-up: mean 26 weeks	The mean maternal confidence/competence mean scores post-treatment – available case analysis (no-risk populations) in the intervention groups was 0.3 standard deviations higher (0.24 lower to 0.84 higher)	54 (1 study)	⊕⊕⊖⊖ low[1,2]	SMD 0.3 (-0.24 to 0.84)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total population size is less than 400 (a threshold rule-of-thumb)
[2] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Mother-infant attachment: mindfulness training versus treatment as usual

There was single study (N=21) low quality evidence for a large benefit of mindfulness training for women in the postnatal period with no identified risk factors on maternal confidence/competence (p=0.002) (Table 108).

Table 108: Summary of findings table for effects of mindfulness training compared with treatment as usual on preventing mother–infant attachment problems for women with no identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	 Corresponding risk Mother-infant attachment: Mindfulness training versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Maternal confidence/competence mean scores post-treatment – available case analysis (no-risk populations) Parental Evaluation Scale: Maternal self-efficacy Follow-up: mean 11 weeks		The mean maternal confidence/competence mean scores post-treatment – available case analysis (no-risk populations) in the intervention groups was 1.59 standard deviations higher (0.56 to 2.62 higher)		21 (1 study)	⊕⊕⊖⊖ low[1]	SMD 1.59 (0.56 to 2.62)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total population size is less than 400 (a threshold rule-of-thumb)

7.4.27 Clinical evidence for preventative effects on poor quality of life outcomes for women with no identified risk factors (by intervention)

Summary of findings can be found in the tables presented in this section. The full GRADE evidence profiles and associated forest plots can be found in Appendix 22 and Appendix 19, respectively.

Quality of life: structured psychological interventions (CBT or IPT) versus treatment as usual

A single study (N=1,299-1,749) found no evidence for clinically significant benefits (despite statistical significance) of CBT for women in the postnatal period with no identified risk factors on maternal stress (p=0.03), impaired life functioning (p=0.07) or wellbeing (p=0.002) (Table 109).

Table 109: Summary of findings table for effects of structured psychological interventions (CBT or IPT) compared with treatment as usual on preventing poor quality of life outcomes for women with no identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Quality of life: Structured psychological interventions (CBT or IPT) versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Parental stress mean scores post-treatment – available case analysis (no-risk populations) PSI Follow-up: mean 26 weeks		The mean parental stress mean scores post-treatment – available case analysis (no-risk populations) in the intervention groups was 0.12 standard deviations higher (0.01 to 0.23 higher)		1,299 (1 study)	⊕⊕⊕⊖ moderate[1]	SMD 0.12 (0.01 to 0.23)
Impaired functioning mean scores post-treatment – available case analysis (no-risk populations) CORE-OM: Life functioning		The mean impaired functioning mean scores post-treatment – available case analysis (no-risk populations) in the intervention groups was 0.09 standard deviations lower		1,747 (1 study)	⊕⊕⊕⊖ moderate[1]	SMD -0.09 (-0.18 to 0.01)

Follow-up: mean 26 weeks	(0.18 lower to 0.01 higher)			
Wellbeing mean scores post-treatment – available case analysis (no-risk populations) CORE-OM: Well-being Follow-up: mean 26 weeks	The mean wellbeing mean scores post-treatment – available case analysis (no-risk populations) in the intervention groups was 0.15 standard deviations lower (0.25 to 0.06 lower)	1,749 (1 study)	⊕⊕⊕⊖ moderate[1]	SMD -0.15 (-0.25 to -0.06)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Paper omits data

Quality of life: listening visits versus treatment as usual

A single study (N=1,407-1,800) found no evidence for clinically significant benefits (despite statistical significance) of listening visits for women in the postnatal period with no identified risk factors on maternal stress (p=0.002), impaired life functioning (p=0.08) or wellbeing (p=0.002) (Table 110).

Table 110: Summary of findings table for effects of listening visits compared with treatment as usual on preventing poor quality of life outcomes for women with no identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Quality of life: Listening visits versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Parental stress mean scores post-treatment – available case analysis (no-risk populations)		The mean parental stress mean scores post-treatment – available case analysis (no-risk populations) in the intervention groups was		1,407 (1 study)	⊕⊕⊕⊖ moderate[1]	SMD 0.17 (0.06 to 0.27)

PSI Follow-up: mean 26 weeks	0.17 standard deviations higher (0.06 to 0.27 higher)			
Impaired functioning mean scores post-treatment – available case analysis (no-risk populations) CORE-OM: Life functioning Follow-up: mean 26 weeks	The mean impaired functioning mean scores post-treatment – available case analysis (no-risk populations) in the intervention groups was 0.08 standard deviations lower (0.18 lower to 0.01 higher)	1,798 (1 study)	⊕⊕⊕⊖ moderate[1]	SMD -0.08 (-0.18 to 0.01)
Wellbeing mean scores post-treatment – available case analysis (no-risk populations) CORE-OM: Well-being Follow-up: mean 26 weeks	The mean wellbeing mean scores post-treatment – available case analysis (no-risk populations) in the intervention groups was 0.15 standard deviations lower (0.24 to 0.05 lower)	1,800 (1 study)	⊕⊕⊕⊖ moderate[1]	SMD -0.15 (-0.24 to -0.05)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Paper omits data

Quality of life: home visits versus treatment as usual

A single study (N=465-513) found no evidence for clinically or statistically significant effects of home visits for women in the postnatal period with no identified risk factors on social support at post-treatment (p=0.87) or at intermediate follow-up (p=0.54) (Table 111).

Table 111: Summary of findings table for effects of home visits compared with treatment as usual on preventing poor quality of life outcomes for women with no identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Quality of life: Home visits versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Social support mean scores post-treatment – available case analysis (no-risk populations) Duke Functional Social Support Follow-up: mean 6 weeks		The mean social support mean scores post-treatment – available case analysis (no-risk populations) in the intervention groups was 0.01 standard deviations higher (0.16 lower to 0.19 higher)		513 (1 study)	⊕⊕⊕⊖ moderate[1]	SMD 0.01 (-0.16 to 0.19)
Social support mean scores Intermediate follow-up (17-24 weeks post-intervention) – available case analysis (no-risk populations) Duke Functional Social Support Follow-up: mean 26 weeks		The mean social support mean scores intermediate follow-up (17-24 weeks post-intervention) – available case analysis (no-risk populations) in the intervention groups was 0.06 standard deviations higher (0.13 lower to 0.24 higher)		465 (1 study)	⊕⊕⊕⊖ moderate[1]	SMD 0.06 (-0.13 to 0.24)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] High risk of selection bias due to statistically significant baseline group differences for incidence of twins, use of TENS during labour, and adults living with the mother

Quality of life: mother-infant relationship intervention versus enhanced treatment as usual

A single study (N=54) found no evidence for a clinically or statistically significant effect of a mother–infant relationship intervention relative to enhanced treatment as usual (monitoring) for women in the postnatal period with no identified risk factors on maternal stress (p=0.14) (Table 112).

Table 112: Summary of findings table for effects of a mother–infant relationship intervention compared with enhanced treatment as usual on preventing poor quality of life outcomes for women with no identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Quality of life: Mother-infant relationship interventions versus Enhanced TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Parental stress mean scores post-treatment – ITT analysis (no-risk populations) Daily Hassles Scale: Intensity Follow-up: mean 26 weeks		The mean parental stress mean scores post-treatment – ITT analysis (no-risk populations) in the intervention groups was 0.4 standard deviations lower (0.94 lower to 0.14 higher)		54 (1 study)	⊕⊕⊖⊖ low[1,2]	SMD -0.4 (-0.94 to 0.14)
Parental stress mean scores post-treatment – available case analysis (no-risk populations) Daily Hassles Scale: Intensity Follow-up: mean 26 weeks		The mean parental stress mean scores post-treatment – available case analysis (no-risk populations) in the intervention groups was 0.4 standard deviations lower (0.94 lower to 0.14 higher)		54 (1 study)	⊕⊕⊖⊖ low[1,2]	SMD -0.4 (-0.94 to 0.14)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total population size is less than 400 (a threshold rule-of-thumb)
[2] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Quality of life: music therapy versus treatment as usual

A single study (N=77) found no evidence for a clinically or statistically significant effect of music therapy relative to treatment as usual for women in the postnatal period with no identified risk factors on maternal stress (p=0.51) (Table 113).

Table 113: Summary of findings table for effects of music therapy compared with treatment as usual on preventing poor quality of life outcomes for women with no identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Quality of life: Music therapy versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Parental stress mean scores post-treatment – available case analysis (no-risk populations) Perceived Stress Scale Follow-up: mean 2 weeks		The mean parental stress mean scores post-treatment – available case analysis (no-risk populations) in the intervention groups was 0.15 standard deviations higher (0.3 lower to 0.6 higher)		77 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	SMD 0.15 (-0.3 to 0.6)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] High risk of selection bias due to unclear allocation concealment and statistically significant group difference at baseline in education (intervention group were more highly educated than control group)
[2] Total population size is less than 400 (a threshold rule-of-thumb)
[3] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Quality of life: mindfulness training versus treatment as usual

A single study (N=21) found low quality evidence for a large benefit of mindfulness training relative to treatment as usual for women in the postnatal period with no identified risk factors on maternal stress (p=0.02) (Table 114).

Table 114: Summary of findings table for effects of mindfulness training compared with treatment as usual on preventing poor quality of life outcomes for women with no identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Quality of life: Mindfulness training versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Parental stress mean scores post-treatment – available case analysis (no-risk populations) DASS-21: Stress Follow-up: mean 11 weeks		The mean parental stress mean scores post-treatment – available case analysis (no-risk populations) in the intervention groups was 1.14 standard deviations lower (2.1 to 0.18 lower)		21 (1 study)	⊕⊕⊖⊖ low[1]	SMD -1.14 (-2.1 to -0.18)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total population size is less than 400 (a threshold rule-of-thumb)

7.4.28 Clinical evidence for preventative effects on poor retention in services and treatment unacceptability for women with no identified risk factors (by intervention)

Summary of findings can be found in the tables presented in this section. The full GRADE evidence profiles and associated forest plots can be found in Appendix 22 and Appendix 19, respectively.

Retention in services and treatment acceptability (using attrition as a proxy measure): structured psychological interventions (CBT or IPT) versus treatment as usual

There was single study evidence (N=2,324) for harms associated with CBT (indicative of poorer retention in services and lower treatment acceptability) for women in the postnatal period with no identified risk factors with higher attrition for women in the intervention group than in the control group (p=0.004) (Table 115).

Table 115: Summary of findings table for effects of structured psychological interventions (CBT or IPT) compared with treatment as usual on preventing poor retention in services or treatment unacceptability for women with no identified risk factors

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Attrition: Structured psychological interventions (CBT or IPT) versus TAU				
Drop-out Incomplete data at endpoint	Study population		RR 1.3 (1.09 to 1.56)	2,324 (1 study)	⊕⊕⊕⊖ moderate1	
	151 per 1000	196 per 1000 (165 to 236)				
	Moderate					
	151 per 1000	196 per 1000 (165 to 236)				

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Paper omits data

Retention in services and treatment acceptability (using attrition as a proxy measure): listening visits versus treatment as usual

A single study (N=2,297) found no clinically or statistically significant effects of listening visits for women in the postnatal period with no identified risk factors on attrition (p=1.00) (Table 116).

Table 116: Summary of findings table for effects of listening visits compared with treatment as usual on preventing poor retention in services or treatment unacceptability for women with no identified risk factors

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Attrition: Listening visits versus TAU				
Drop-out Incomplete data at endpoint Follow-up: mean 26 weeks	Study population		RR 1 (0.82 to 1.21)	2,297 (1 study)	⊕⊕⊕⊖ moderate1	
	151 per 1000	151 per 1000 (124 to 183)				
	Moderate					
	151 per 1000	151 per 1000 (124 to 183)				

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Paper omits data

Retention in services and treatment acceptability (using attrition as a proxy measure): Psychologically (CBT/IPT)-informed psychoeducation versus enhanced treatment as usual

A single study (N=540) found no evidence for clinically or statistically-significant effects of a psychologically (CBT/IPT)-informed psychoeducational intervention relative to enhanced treatment as usual (non-mental health-focused education and support [booklet and telephone call]) for women in the postnatal period with no identified risk factors on attrition (p=0.74) (Table 117).

Table 117: Summary of findings table for effects of psychologically (CBT/IPT)-informed psychoeducation compared with enhanced treatment as usual on

preventing poor retention in services or treatment unacceptability for women with no identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Attrition: Psychologically (CBT/IPT)-informed psychoeducation versus Enhanced TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Drop-out Incomplete data at endpoint Follow-up: mean 4 weeks	Study population		RR 1.11 (0.61 to 2.01)	540 (1 study)	⊕⊕⊕⊖ moderate1	
	70 per 1000	78 per 1000 (43 to 141)				
	Moderate					
	70 per 1000	78 per 1000 (43 to 141)				

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Retention in services and treatment acceptability (using attrition as a proxy measure): home visits versus treatment as usual

A single study (N=623) found very low quality evidence for moderate benefits of home visits relative to treatment as usual for women in the postnatal period with no identified risk factors on preventing poor retention in services and treatment unacceptability, using attrition as a proxy (p=0.08) (Table 118).

Table 118: Summary of findings table for effects of home visits compared with treatment as usual on preventing poor retention in services or treatment unacceptability for women with no identified risk factors

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Attrition: Home visits versus TAU				
Drop-out Incomplete data at endpoint Follow-up: mean 6 weeks	Study population		RR 0.68 (0.43 to 1.05)	623 (1 study)	⊕⊖⊖⊖ very low1,2,3	
	138 per 1000	94 per 1000 (59 to 145)				
	Moderate					
	138 per 1000	94 per 1000 (59 to 145)				

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 High risk of selection bias due to statistically significant baseline group differences for incidence of twins, use of TENS during labour, and adults living with the mother
2 Total number of events is less than 300 (a threshold rule-of-thumb)
3 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Retention in services and treatment acceptability (using attrition as a proxy measure): mindfulness training versus treatment as usual

There was single study evidence (N=26) for harms associated with mindfulness training (indicative of poorer retention in services and lower treatment acceptability) for women in the postnatal period with no identified risk factors with higher attrition for women in the intervention group than for women who received treatment as usual (p=0.09). However, confidence in this effect estimate was low due to very serious imprecision (Table 119).

Table 119: Summary of findings table for effects of mindfulness training compared with treatment as usual on preventing poor retention in services or treatment unacceptability for women with no identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	 Corresponding risk Attrition: Mindfulness training versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Drop-out Incomplete data at endpoint Follow-up: mean 11 weeks	Study population		RR 11 (0.67 to 180.65)	26 (1 study)	⊕⊕⊖⊖ low1,2	
	0 per 1000	0 per 1000 (0 to 0)				
	Moderate					
	0 per 1000	0 per 1000 (0 to 0)				

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total number of events is less than 300 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

7.4.29 Health economic evidence

Systematic literature review

The systematic literature search identified two eligible UK studies (Barlow et al., 2007 and McIntosh et al., 2009; Petrou et al., 2006), one study conducted in Chile (Aracena et al., 2009) and one in Australia (Hiscock et al., 2007) that assessed prevention interventions for developing mental health problems in pregnancy or the postnatal period. Details on the methods used for the systematic search of the economic literature are described in Chapter 3. References to included studies and evidence tables for all economic studies included in the guideline systematic literature review are presented in Appendix 21. Completed methodology checklists of the studies are provided in Appendix 20. Economic evidence profiles of studies considered during guideline development (that is, studies that fully or partly met

the applicability and quality criteria) are presented in Appendix 22, accompanying the respective GRADE clinical evidence profiles.

Barlow and colleagues (2007) evaluated the cost effectiveness of a home visiting programme compared with standard care in vulnerable pregnant women. Women were screened using a range of demographic and socioeconomic criteria (for example, presence of mental health problems or housing problem). The programme involved health visitors trained in the Nurse-Family Partnership Model providing intensive weekly home visiting services from 6 months antenatally to 12 months after childbirth. Standard care was defined as locally available services. This was an economic evaluation undertaken alongside an RCT (BARLOW2007) (n=131) conducted in the UK. The study by McIntosh and colleagues (2009) is based on the same RCT but reports additional analyses. The main analysis was conducted from a public sector perspective plus informal care but authors conducted sensitivity analyses considering a healthcare perspective. The study considered a range of direct healthcare costs including primary and secondary care, direct non-healthcare costs (that is, social worker, alcohol/drug support, child and family team, foster care, adoption services, family centre, Sure Start, Home Start); also the costs accruing to Housing department, legal advice centre, Citizens Advice Bureau, court and to the police; and childcare costs (that is, crèche, playgroup and private childcare). The resource use estimates were based on the RCT and other published sources. The unit costs were obtained from local and national sources. The measure of outcome for the economic analysis was the proportion of infants identified as being ill-treated on the basis of child protection proceedings between 6 and 12 months after childbirth, improvement in maternal sensitivity and infant cooperativeness components of CARE-Index scores; and time of infant exposure to abuse and neglect. The CARE-Index is a measure that assesses mother–infant interaction from birth to about two years of age based on a short, videotaped play interaction of 3-5 minutes. The measure assesses mothers on three scales: sensitivity, control and unresponsiveness. There are also four scales for infants: cooperativeness, compulsivity, difficultness, and passivity. The time horizon of the main analysis was 18 months, however when using the time of infant exposure to abuse and neglect as an outcome of the economic analysis costs were modelled for 5 years. The authors assumed that exposure to abuse and neglect would continue throughout the preschool period, and that the neglect would be identified as soon as the child went to school at the age of 5 years (for example, assuming that neglect was identified when the child was 6 months old, intervention would have prevented 4.5 years of abuse and neglect); the costs considered over this period of time included foster care and adoption costs.

The intervention resulted in a greater proportion of infants being identified as ill-treated between 6 and 12 months compared with standard care, (0.059 versus 0.000, respectively; difference 0.059, *p* value was non-significant); improvement in maternal sensitivity component of CARE-Index score: 9.27 versus 8.20 for intervention and standard care, respectively (difference of 1.07 points); improvement in infant cooperativeness component of CARE-Index score: 9.35 and 7.92 for intervention and standard care, respectively (difference of 1.43 points). For a reduction in time of

exposure to abuse the difference was 1.9 months in favour of the intervention. From a public sector perspective (and informal care) the mean total costs per mother–infant dyad over 18 months were £7,120 for the intervention and £3,874 for standard care, a difference of £3,246 ($p < 0.05$) in 2003/04 prices. Similarly, when considering only health service costs, the mean total costs per mother–infant dyad over 18 months were £5,685 for intervention and £3,324 for standard care, a difference of £2,360 ($p < 0.05$).

From a public sector perspective (and informal care) the cost per extra infant identified as being ill-treated was £55,016; per extra unit of improvement on maternal sensitivity and infant cooperativeness components of CARE-Index it was £2,723 and £2,023, respectively; and £1,691 per additional month reduced of infant exposure to abuse and neglect. From a healthcare perspective the cost per extra infant identified as being ill-treated was £40,000; per extra unit of improvement on maternal sensitivity and infant cooperativeness components of CARE-Index it was £2,178 and £1,621, respectively; and £1,229 for a reduction in infant exposure to abuse and neglect by one month.

From a public sector perspective (and informal care) probabilistic analysis indicated that at a willingness-to-pay (WTP) of £16,100 per unit improvement on the maternal sensitivity component of CARE-Index the probability that the intervention is cost effective was 0.95 and at WTP of £4,000 per unit improvement on infant cooperativeness component of CARE-Index the probability that the intervention is cost effective was 0.95. Moreover, at WTP of £1,400 for a reduction in infant exposure to abuse and neglect by one month the probability that the intervention is cost effective was 0.75 and at WTP £3,100 this probability increased to 0.95. From a healthcare perspective when WTP is £13,900 and £2,700 per unit improvement on maternal sensitivity component of CARE-Index and on infant cooperativeness component of CARE-Index, respectively, the probability that intervention is cost effective was 0.95. Deterministic sensitivity analyses were very limited and were conducted only on the ICER estimated from a public sector perspective plus informal care. It was found that ranging the proportion of infants identified as being ill-treated from 0.03 to 0.13 (base-case 0.06), the cost for a reduction in infant exposure to abuse and neglect by one month ranged from £2,505 to £1,284. Overall results suggest that intervention provides better outcomes however at an additional cost.

The analysis was judged by the GDG to be partially applicable to this guideline review and the NICE reference case. In the base case analysis the authors explored the cost effectiveness from a public sector perspective (plus informal care). Moreover, the authors did not attempt to estimate QALYs which made it difficult to interpret the cost-effectiveness results and to compare the findings with other studies. Also, the sensitivity analysis was very limited. However, overall, given the

data limitations in this area, this was a well conducted study and was judged by the GDG to have only minor methodological limitations.

Petrou and colleagues (2006) evaluated the cost effectiveness of listening visits compared with standard care. Standard care was defined as care provided by local primary care teams. The intervention entailed research therapists visiting women in their homes at 35 and 37 weeks antenatally; on days 3, 7, and 17 after childbirth, and then weekly up to 8 weeks. Study population comprised women at high risk of developing depression in the postnatal period [women who scored ≥ 24 on the predictive index developed by Cooper and colleagues (1996) at 26-28 weeks of gestation]. This was an economic evaluation undertaken alongside an RCT (n=151) conducted in the UK. The time horizon of the analysis was 18 months; healthcare and informal care costs were considered. The study estimated a range of costs including community care, day care, hospital outpatient and inpatient care, paediatric care, child care and home help. The authors did not report healthcare costs separately, consequently it was not possible to estimate costs from the NHS and PSS perspective. The resource use estimates were based on the RCT (n=151) and the unit costs were obtained from local and national sources. The measure of outcome for the economic analysis was the number of months in depression in the postnatal period. In the analysis, costs and health effects beyond 12 months were discounted at an annual rate of 6% and 1.5%, respectively.

At 18 months the intervention resulted in fewer months of depression in the postnatal period per woman, 2.21 months versus 2.70 months, difference of -0.49 months (p=0.41). The mean cost per mother–infant dyad over 18 months was £2,397 for the intervention and £2,278 for standard care in 2000 prices, difference of £120 (p=0.72). The cost per month in depression avoided was estimated to be £244. The authors also conducted a range of sensitivity analyses. According to the deterministic sensitivity analysis when varying community service utilisation from 10 to 30% the ICER ranged from £422 to £780; when increasing or decreasing per diem cost for inpatient care by 20% the ICER ranged from £41 to £446; when ranging the discount rate for costs and health effects from 0% to 10% the ICER ranged from £351 to £198; and when setting discount rate for costs and health effects at 3% the ICER increased to £302 per month of depression avoided. Probabilistic analysis indicated that at WTP of £1,000 and £2,000 per month of depression avoided the probability of the intervention being cost effective was 0.71 and 0.77, respectively. Results suggest that intervention provides better outcome at an additional cost, although the differences in costs and clinical outcomes were not statistically significant.

The analysis was judged by the GDG to be partially applicable to this guideline review and the NICE reference case. The authors included some cost categories that are not relevant to the NHS and PSS perspective (that is, informal care) and some of the unit costs were derived from local sources which may limit the generalisability of the findings. Also, NICE recommends discounting both costs and health effects at an annual rate of 3.5%, but in the analysis a discount rate of 6% and 1.5% was used for

costs and health effects, respectively. Nevertheless, as indicated by the sensitivity analysis the discount rate had a minimal effect on the ICER. The estimate of relative treatment effect was obtained from a single RCT and the authors have not attempted to estimate QALYs, which made it difficult to interpret the cost-effectiveness results and to compare the findings with other studies. Overall this was a well conducted study and was judged by the GDG to have only minor methodological limitations.

Aracena and colleagues (2009) evaluated the cost effectiveness of home visiting service compared with standard care in Chile. The intervention involved home visits from health educators, starting in the third trimester of pregnancy and continuing until the child reached 1 year of age; in total, the women had 12 home visits of 1 hour each throughout the year. Standard care was defined as standard prenatal and well-baby care at local health centres and consisted of ten prenatal consultations with midwife at the local health centres. The study population comprised of young women from low socioeconomic backgrounds who conceived their first child when they were between 14-19 years old. This was an economic evaluation undertaken alongside an RCT (ARACENA2009) (n=90). The time horizon of the analysis was 15 months and the perspective of the healthcare payer was adopted. The study estimated healthcare, administrative and logistical costs. The resource use estimates were based on registries of health centres, and the source of unit costs was not specified. The measure of outcome for the economic analysis was an improvement in the Goldberg's Depression Scale score. Neither costs nor health effects were discounted in the economic analysis, but such discounting was not necessary because the time horizon was 15 months.

Over 15 months the intervention resulted in an improvement in Goldberg's Depression Scale score: 10.94 (SD 5.85) versus 13.85 (SD 6.99), intervention and standard care groups, respectively (difference of -2.91 points, p=0.031). The costs in the study were measured in US dollars and the cost year was not reported. The median cost per mother–infant dyad at 15 months was $90 for the intervention and $50 for the standard care group, showing a difference of $40. The cost per additional score reduction on the Goldberg's Depression Scale was estimated to be $13.50. Results suggest that home visits provide the better outcome; however, this comes at an additional cost.

The GDG considered the analysis to be partially applicable to this guideline review and the NICE reference case. The study was conducted in Chile and the type of healthcare costs considered in the analysis are unclear. Moreover, the authors did not attempt to estimate QALYs, which made it difficult to interpret the cost-effectiveness results and to compare the findings with other studies. The estimate of relative treatment effect was obtained from a single RCT, the resource use estimates were derived from registries of local health centres which may limit the generalisability of the findings to the UK setting; and the source of unit costs was unclear. Also, statistical analysis was done only for outcomes and not for costs. As a result, this study was judged by the GDG to have potentially serious methodological limitations.

Hiscock and colleagues (2007) evaluated the cost effectiveness of an infant sleep training intervention compared with standard care. This was an economic evaluation undertaken alongside an RCT (HISCOCK2002 [Hiscock & Wake, 2002]) (n=328) conducted in Australia. Infant sleep intervention entailed mothers attending three consultations at their local maternal and child health centres. Mothers were given a choice of two behavioural interventions: (1) 'controlled crying' whereby parents respond to their infant's cry at increasing time intervals, to allow independent settling or (2) 'camping out' sitting with the infant until they fall asleep and gradually removing parental presence over 3 weeks. In standard care group mothers were given an infant sleep leaflet only. The study population comprised mothers of 4-month-old infants attending a consultation at a maternal and child health centre and reporting an infant sleep problem. The time horizon of the analysis was 12 months; costs included healthcare and informal care. The study included costs associated with consultations for sleep advice at maternal and child health centres, non-maternal and child health nurse professional healthcare (such as parenting centres and family doctor), non-professional care (such as books, care provided by relatives), intervention, and nurse training programme. The resource use estimates were based on the RCT (n=309) and the unit costs were obtained from local and national sources. The measure of outcome for the economic analysis was maternal report of infant sleep problem; presence of depression symptoms (measured using EPDS); and SF-12 mental health domain scores.

The intervention resulted in fewer mothers reporting an infant sleep problem: 39% and 55% in intervention and standard care groups, respectively (difference of -16%, p=0.004). The intervention also resulted in a reduction in EPDS scores: 5.9 and 7.2 in intervention and standard care groups, respectively (difference of -1.7 points, p=0.001); and improvement in SF-12 mental health domain scores: 49.7 and 46.1 in intervention and standard care groups, respectively (difference of 3.9 points, p <0.001). The costs in the study were measured in British pounds, expressed in 2007 prices. The mean cost per family over 12 months was £97 (SD £249) for the intervention and £117 (SD £330) for standard care, respectively, difference of -£19.44 (p=0.55). Results suggest that intervention provides better outcomes at a slightly lower cost, and thus is a dominant intervention.

The analysis was judged by the GDG to be partially applicable to this guideline review and the NICE reference case. This study was conducted in Australia where the healthcare system is sufficiently similar to the UK NHS. However, the analysis included cost categories beyond the NHS and PSS perspective (that is, costs associated with informal care). Also, the authors did not attempt to estimate QALYs but this did not affect interpretation of the results, since intervention was found to be dominant. Also, the source of unit costs was unclear. Overall, the study was judged by the GDG to have only minor methodological limitations.

Overall conclusions from existing economic evidence

The existing economic evidence on psychological and psychosocial interventions for the prevention of mental health problems in pregnancy or postnatal period is very limited. The systematic literature review identified two UK-based studies and two non-UK studies. None of the studies were directly applicable to the NICE decision-making context. Both UK-based studies found prevention interventions (home visiting and listening visits) to result in better outcomes however at an additional cost. This finding is supported by evidence from studies conducted in Chile where home visiting resulted in better outcomes but also led to an increase in costs. In an Australian study an infant sleep training intervention resulted in better outcomes at a slightly lower cost, and thus was found to be a dominant intervention. The results from these studies are not easy to interpret due to lack of use of QALYs as a measure of outcome in the majority of the studies, and difficulty in judging whether the additional cost per non-QALY outcomes such as a month in depression avoided, point improvement on a depression scale or point change on mother infant interaction scales represent good value for money. Overall, the results are inconclusive, as they do not use QALYs and it is difficult to judge whether the reported extra benefits associated with the prevention interventions are worth the extra costs associated with their provision.

7.5 PSYCHOLOGICAL AND PSYCHOSOCIAL INTERVENTIONS FOR THE TREATMENT OF MENTAL HEALTH PROBLEMS

7.5.1 Introduction

Despite the evidence illustrating that mental health problems are common, debilitating and have a broader direct effect on the woman's fetus and newborn infant, and that medication is less acceptable in pregnancy and the postnatal period than at other times, the efficacy and acceptability of psychological or psychosocial treatments in pregnancy and the postnatal period has not been extensively researched. Historically, there has been an emphasis on postnatal depression and most treatment research has been carried out in this field. Treatment in pregnancy and the period has been aimed at preventing the development of postnatal mental health problems, making such studies difficult to interpret.

There seem to be widely held but poorly substantiated beliefs that neither pregnancy nor the early postnatal period are times to make life changes and that psychological or psychosocial treatment may be harmful and should be avoided. This, in combination with the fact that being pregnant or having a newborn infant clearly leads to difficulties in accessing standard psychological treatments in general services that may have long waiting lists and inflexible clinic times, has exacerbated the problems of access to psychological treatments for this group. A number of attempts have been made to modify psychological treatments for pregnancy and the postnatal period, involving a broad range of healthcare professionals delivering

treatments at home or in groups. Research comparing these modified treatments with standardised therapies such as CBT and IPT has not been undertaken and the advantage in the modification remains unclear.

7.5.2 Clinical review protocol (treatment)

The review protocol summary, including the review question(s) and the eligibility criteria used for this section of the guideline, can be found in Table 120. A complete list of review questions can be found in Appendix 8; further information about the search strategy can be found in Appendix 10; the full review protocols can be found in Appendix 9.

The review strategy was to evaluate the clinical effectiveness of the interventions using meta-analysis. However, in the absence of adequate data, the available evidence was synthesised using narrative methods. An analysis of all interventions was conducted and graded. Following this, sub-analysis was conducted (dependent on available data), based on baseline diagnostic status (clinical diagnosis [usually assessed using structured psychiatric interview]; symptoms [above a pre-specified threshold on a rating scale]; subthreshold symptoms [just below a pre-specified threshold on a rating scale]), treatment timing, mode of delivery, format (individual and/or group), and intensity. Where possible both an available case analysis and an intention-to-treat (ITT) analysis (WCS) were used.

Table 120: Clinical review protocol summary for the review of psychological and psychosocial interventions for the treatment of mental health problems

Component	Description
Review question(s)	RQ 1.10 For women with mental disorders who are pregnant or in the postnatal period, what are the benefits and/or potential harms of **psychosocial** interventions to treat mental health problems? RQ 1.14 For women with mental disorders who are pregnant or in the postnatal, what are the benefits and/or potential harms of interventions targeted at improving the **quality of the mother-child interaction?** RQ 1.15 What is the role of the family, carers and peers in the treatment and support of women with mental health disorders in pregnancy and the postnatal period?
Population	**Included** Women who have mental health problems during pregnancy and the postnatal period (from delivery to the end of the first year). Include:- • Women with subthreshold symptoms (but no formal diagnosis of a mental health problem) • Women with a formal diagnosis of mild, moderate and severe disorders **Exclude women:-** who are not pregnant or in the postnatal period (up to 1 year postnatal)
Intervention(s)	Psychological or psychosocial interventions, including: • Home visits

	• Listening visits (non-directive counselling) • Mother-infant relationship interventions • Peer-mediated support and support groups • Post-miscarriage interventions • Post-traumatic birth counselling • Pre-delivery discussion and psychoeducation (for tokophobia) • Protocols for women following stillbirth • Psychologically (CBT or IPT)-informed psychoeducation • Self-help and facilitated self-help • Structured psychological interventions (CBT or IPT)
Comparison	Treatment as usual, enhanced treatment as usual, no treatment, wait-list control, other active interventions
Critical outcomes	**Maternal Outcomes** • **Symptom-based** o Diagnosis of mental disorder o Symptomatology o Relapse o Use of drugs/alcohol • **Service utilisation** o Hospitalisation o Retention in services (assessed through drop-out rates as a proxy measure) o Health service utilisation (for instance, use of psychiatric services) • **Experience of care** o Satisfaction (validated measures only, specific items will not be analysed) o Acceptability of treatment (assessed through questioning or through including drop-out as a proxy measure) • **Quality of life** o Quality of life measures o Functional disability o Social functioning o Social support o Self-esteem o Perceived parenting stress o Maternal confidence o Preservation of rights • **Harm** o Side effects (including drop-out because of side effects) o Maternal mortality and serious morbidity including self-harm and suicide attempts • **Quality of mother–infant interaction** o Quality of mother–infant interaction (including maternal sensitivity and child responsivity) o Maternal attitude towards motherhood o Establishing or continuing breastfeeding **Infant outcomes (no restriction on length of follow-up)** • Fetal and infant physical development (including congenital malformations)

	• Side effects (especially of pharmacological interventions for the fetus and for the infant if breastfeeding) • Apgar score • Birth weight • Admission to neonatal intensive care unit • Cognitive development of the infant • Emotional development of the infant • Physical development of the infant • Prevention of neglect or abuse of the infant • Optimal care of infant (for example, vaccinations, well-baby check-ups) • Fetal/infant mortality • Fetal/infant morbidity • Service use o Planned (health visitor, vaccinations, well-baby check-ups) o Unplanned (emergency department visits, inpatient, urgent or acute care) o Social service involvement
Study design	Systematic reviews of RCTs Primary RCTs For protocols for women following stillbirth, cohort studies were included
Note.	

7.5.3 Studies considered (treatment)

Seventy-four RCTs reported across 93 papers met the eligibility criteria for this review: AMMERMAN2013A/2013B, ARMSTRONG1999/2000/FRASER2000 (Armstrong et al., 1999; Armstrong et al., 2000; Fraser et al., 2000), ARMSTRONG2003 (Armstrong & Edwards, 2003), ARMSTRONG2004 (Armstrong & Edwards, 2004), AUSTIN2008 (Austin et al., 2008), BERNARD2011 (Bernard et al., 2011), BILSZTA2012 (Bilszta et al., 2012), BURNS2013/PEARSON2013B, CHEN2000 (Chen et al., 2000), CHO2008 (Cho et al., 2008), COOPER2003/MURRAY2003, DENNIS2003 (Dennis, 2003), DENNIS2009 (Dennis et al., 2009), DUGGAN2007/CALDERA2007 (Duggan et al., 2007; Caldera et al., 2007), DUGRAVIER2013/GUEDENEY2013 (Dugravier et al., 2013; Guedeney et al., 2013), ELMOHANDES2008 (El-Mohandes et al., 2008), FIELD2013A (Field et al., 2013a), GAMBLE2005 (Gamble et al., 2005), GAO2010/2012 (Gao et al., 2010; Gao et al., 2012), GUARDINO2014 (Guardino et al., 2014), GROTE2009, HAGAN2004 (Hagan et al., 2004), HAYDEN2012 (Hayden et al., 2012), HISCOCK2002, HISCOCK2007/2008 (Hiscock et al., 2007; Hiscock et al., 2008), HOLDEN1989 (Holden et al., 1989), HONEY2002 (Honey et al., 2002), HOROWITZ2001 (Horowitz et al., 2001), KAAYA2013 (Kaaya et al., 2013), KERSTING2011 (Kersting et al., 2011), KOZINSZKY2012 (Kozinszky et al., 2012), LE2011 (Le et al., 2011), LETOURNEAU2011 (Letourneau et al., 2011), LEUNG2012 (Leung & Lam, 2012), MILGROM2005B (Milgrom et al., 2005b), MILGROM2011A (Milgrom et al., 2011a), MILGROM2011B (Milgrom et al., 2011b), MISRI2000 (Misri et al., 2000), MORRELL2009A/2009B/2011/BRUGHA2011 (Morrell et al., 2009a; Morrell et al.,

2009b; Morrell et al., 2011; Brugha et al., 2011), MULCAHY2010 (Mulcahy et al., 2010), MUNOZ2007/URIZAR2011 (Muñoz et al., 2007; Urizar & Muñoz, 2011), NEUGEBAUER2006 (Neugebauer et al., 2006), NIKCEVIC2007 (Nikčević et al., 2007), OHARA2000, OMAHEN2013A (O'Mahen et al., 2013a), OMAHEN2013B (O'Mahen et al., 2013b), OMAHEN2013C (O'Mahen et al., 2013c), ORTIZCOLLADO2014 (Ortiz-Collado et al., 2014), PINHEIRO2014 (Pinheiro et al., 2014), PRENDERGAST2001 (Prendergast & Austin, 2001), RAHMAN2008, ROMAN2009 (Roman et al., 2009), ROUHE2012/SALMELAARO2012 (Rouhe et al., 2012; Salmela-Aro et al., 2012); SAISTO2001 (Saisto et al., 2001), SALOMONSSON2011 (Salomonsson & Sandell, 2011), SILVERSTEIN2011 (Silverstein et al., 2011), SIMAVLI2014 (Simavli et al., 2014), SLEED2013 (Sleed et al., 2013), SPINELLI2003 (Spinelli & Endicott, 2003), STEIN2006 (Stein et al., 2006), SWANSON2009 (Swanson et al., 2009), TAMAKI2008 (Tamaki, 2008), TANDON2011/2014/MENDELSON2013 (Tandon et al., 2011; Tandon et al., 2014; Mendelson et al., 2013), TIMPANO2011 (Timpano et al., 2011), VANDOESUM2008/KERSTENALVAREZ2010 (van Doesum et al., 2008; Kersten-Alvarez et al., 2010), VIETEN2008 (Vieten & Astin, 2008), WEIDNER2010 (Weidner et al., 2010), WICKBERG1996 (Wickberg & Hwang, 1996), WIGGINS2005 (Wiggins et al., 2005), WIKLUND2010 (Wiklund et al., 2010), ZELKOWITZ2008/2011/ FEELEY2012 (Zelkowitz et al., 2008; Zelkowitz et al., 2011; Feeley et al., 2012), ZLOTNICK2001 (Zlotnick et al., 2001), ZLOTNICK2006 (Zlotnick et al., 2006), ZLOTNICK2011 (Zlotnick et al., 2011). All of these studies were published in peer-reviewed journals between 1989 and 2014. In addition, 20 studies were excluded from the review. The most common reasons for exclusion were that data could not be extracted, the intervention was outside the scope (organisation of care), non-randomised group allocation, or the paper did not report mental health outcomes. Further information about both included and excluded studies can be found in Appendix 18.

Of the 74 included RCTs, there were 14 studies (N=2,099) involving a comparison of structured psychological interventions (CBT or IPT) and treatment as usual or enhanced treatment as usual, two studies (N=438) compared CBT to listening visits, one study (N=60) compared CBT and relational constructivist therapy, and one study (N=48) involved a comparison of IPT and a support group (Table 121).

Three RCTs (N=1,136) involved a comparison of facilitated self-help and treatment as usual, and two studies involved a comparison of post-miscarriage self-help and treatment as usual (N=255), one study compared post-miscarriage facilitated self-help with treatment as usual (N=171; Table 122).

Five studies (N=1,018) compared listening visits (non-directive counselling) and treatment as usual, one study (N=146) involved a comparison of directive counselling and treatment as usual, three studies (N=269) compared post-miscarriage counselling and treatment as usual or enhanced treatment as usual, and one study (N=103) compared post-traumatic birth counselling and treatment as usual (Table 123).

Four studies (N=867) involved a comparison of social support (peer-mediated support or support group) and treatment as usual, 16 studies (N=2,955) compared psychologically (CBT/IPT)-informed psychoeducation and treatment as usual or enhanced treatment as usual, one study (N=38) involved a comparison between IPT-informed psychoeducation and a non-mental health-focused education and support group, one study (N=331) compared non-mental health-focused education and support (group counselling intervention for HIV-positive women) and treatment as usual, five studies (N=1,616) compared home visits with treatment as usual or enhanced treatment as usual, and two studies (N=547) compared pre-delivery discussion/psychoeducation for tokophobia and treatment as usual (Table 124). Six studies (N=691) compared mother–infant relationship interventions and treatment as usual, one study (N=51) involved a comparison of mother–infant relationship intervention with video feedback and mother–infant relationship intervention with verbal feedback (this trial also included a TAU arm but this data could not be extracted due to non-random assignment to that condition), one study (N=80) compared mother–infant relationship intervention and listening visits (participants in both conditions also received facilitated self-help aimed at their eating disorder), and one study (N=29) compared a co-parenting intervention and enhanced treatment as usual (Table 125).

Two studies (N=394) involved a comparison of infant sleep training (controlled crying) and treatment as usual or enhanced treatment as usual, one study (N=161) compared music therapy during birth and treatment as usual, two studies (N=276) compared a psychosomatic intervention and treatment as usual, and two studies (N=81) compared mindfulness training and treatment as usual or enhanced treatment as usual (Table 126).

Finally, there was one study (N=20) that compared a combined psychosocial (informal support group) and physical (exercise) with enhanced treatment as usual, and one study (N=24) that involved a comparison of social support and physical exercise (Table 127).

For the review of psychosocial treatment for alcohol or substance misuse, three Cochrane reviews met the eligibility criteria for this review: STADE2009B (Stade et al., 2009b), TERPLAN2007 (Terplan & Liu, 2007), TURNBULL2012 (Turnball & Osborn, 2012). In addition, five individual studies (MARAIS2011 [Marais et al., 2011], OSTERMAN2012 [Osterman & Dyehouse, 2012], OSTERMAN2014 [Osterman et al., 2014], WINHUSEN2008 [Winhusen et al., 2008], YONKERS2012 [Yonkers et al., 2012] met the eligibility criteria for this review and were used to update the Cochrane reviews. An additional three primary RCTs (FLEMING2008 [Fleming et al., 2008], ONDERSMA2014 [Ondersma et al., 2014], SILVERMAN2002 [Silverman et al., 2002]) met eligibility criteria for this review but not for any of the Cochrane reviews and were analysed separately (Table 128). An additional Cochrane review was identified by the search, however, no suitable trials were identified by this

review and as a result there was no data that could be extracted (LUI2008 [Lui et al., 2008]). A further seven studies were identified by the search for this review (and were not reviewed in any of the Cochrane reviews) but were excluded on the following basis: systematic review with no new data (Gilinsky et al., 2011); no mental health outcome reported (Armstrong et al., 2009); data could not be extracted (Kropp et al., 2010; Ondersma et al., 2012); intervention was delivered greater than 1 year into the postnatal period (Suchman et al., 2010; Suchman et al., 2011; Suchman et al., 2012).

Table 121: Study information table for trials included in the meta-analysis of structured psychological interventions (CBT or IPT) versus any alternative management strategy

	Structured psychological interventions (CBT or IPT) versus TAU or Enhanced TAU	**CBT versus Listening visits**	**CBT versus relational constructivist therapy**	**IPT versus Support group**
Total number of trials (number of participants)	14 (2,099)	2 (438)	1 (60)	1 (48)
Study ID	(1) AMMERMAN2013A/2013B (2) BURNS2013/PEARSON2013B (3) CHO2008 (4) COOPER2003/MURRAY2003[3] (5) GROTE2009 (6) MILGROM2005B[4] (7) MILGROM2011B (8) MORRELL2009A/2009B/2011/ BRUGHA2011[5] (9) MULCAHY2010 (10) OHARA2000 (11) OMAHEN2013B (12) PRENDERGAST2001 (13) RAHMAN2008 (14) WIKLUND2010	(1) HAYDEN2012 (2) MORRELL2009A/ 2009B/2011/ BRUGHA2011[2]	PINHEIRO2014	FIELD2013A
Country	(1) US (2) UK (3) Korea (4) UK (5) US (6)-(7) Australia (8) UK (9) Australia (10) US (11) UK	(1) US (2) UK	Brazil	US

	(12) Australia (13) Pakistan (14) Sweden			
Mean age of participants (years)	(1) 21.9 (2) 29.2 (3) 29 (4) 27.7 (5) 24.5 (6) 29.7 (7) 31.5 (8) 30.9 (9) 32.2 (10) 29.6 (11) 27 (12) 32.2 (13) 26.7 (14) NR	(1) 31 (2) 30.9	27	24.9
Baseline diagnostic status	(1) Diagnosis of MDD (SCID for DSM-IV) (2) Diagnosis of depression (CIS-R for ICD-10) (3) Diagnosis of depressive disorder (SCID for DSM-IV) (4) Diagnosis of MDD (SCID for DSM-III-R) (5) Diagnosis of depression (SCID for DSM-IV): 85% MDD; 13% dysthymia; 13% comborbid MDD and dysthymia; 6% minor depression (6) Diagnosis of minor depression or MDD (Composite International Diagnosis Interview [CIDI] for DSM-IV) (7) Symptoms of depression (EPDS≥13) (8) Symptoms of depression (EPDS≥12) (9) Diagnosis of MDD (MCMI-III for DSM-IV)	(1) Diagnosis of MDD (Diagnostic Interview Schedule for DSM-IV) (2) Symptoms of depression (EPDS≥12)	Symptoms of depression (BDI≥12)	Diagnosis of MDD or dysthymia (SCID for DSM-IV)

	(10) Diagnosis of major depressive episode (SCID for DSM-IV) (11) Diagnosis of MDD (SCID for DSM-IV) (12) Diagnosis of minor depression or MDD (psychiatric clinical interview for DSM-IV) (13) Diagnosis of major depressive episode (SCID for DSM-IV) (14) Symptoms of depression (EPDS≥12)			
Timing of intervention	(1) Postnatal (2)-(3) Antenatal (4) Postnatal (5) Antenatal and postnatal (6)-(10) Postnatal (11) Antenatal and postnatal (12) Postnatal (13) Antenatal and postnatal (14) Postnatal	(1) Antenatal (2) Postnatal	Postnatal	Antenatal
Mode of delivery	(1)-(14) Face-to-face	(1)-(2) Face-to-face	Face-to-face	Face-to-face
Format	(1)-(5) Individual (6) Group (7)-(8) Individual (9) Individual and group (10)-(14) Individual	(1)-(2) Individual	Individual	Group
Intensity (number of sessions)[1]	(1)-(4) Moderate (9-12 sessions) (5) High (15-21 sessions [including maintenance sessions]) (6) Moderate (11 sessions) (7) Low (4-5 sessions) (8)-(11) Moderate (8-12 sessions) (12) Low (6 sessions) (13) High (16 sessions) (14) Low (3 sessions)	(1)-(2) Moderate (8-10 sessions)	Low (7 sessions)	Moderate (12 sessions)
Length of intervention (weeks)	(1) 15 (2) 12 (3) 18	(1) 10 (2) 8	NR	12

	(4) 10 (5) 44 (6) 12 (7) 6 (8)-(9) 8 (10) 12 (11) NR (12) 6 (13) 48 (14) 3			
Time points[2]	(1) Post-treatment; Short follow-up (2) Post-treatment; Intermediate follow-up (3) Post-treatment (4) Post-treatment; Intermediate follow-up; Long follow-up; Very long follow-up (5) Post-treatment (6) Post-treatment; Long follow-up (7) Post-treatment (8) First measurement (9) Post-treatment; Short follow-up (10) Post-treatment (11) Post-treatment; Short follow-up (12) Post-treatment; Long follow-up (13)-(14) Post-treatment	(1)-(2) First measurement	Post-treatment	Post-treatment
Setting	(1)-(2) Home (3) NR (4) Home (5)-(6) Clinic (primary) (7) Clinic (primary) or hospital (8) Home (9)-(10) NR (11)-(13) Home (14) NR	(1) NR (2) Home	Clinic (secondary)	NR
Intervention	(1) CBT (+ home visits) (2)-(3) CBT (4) IPT (Psychodynamic therapy)	(1)-(2) CBT	CBT	IPT

	(5) IPT (6) CBT (7) CBT ([nurse-led and psychologist-led combined] + GP training) (8) CBT (9)-(10) IPT (11)-(14) CBT			
Comparison	(1) Home visits (2) TAU (3) Enhanced TAU (single session psychoeducation) (4) TAU (5) Enhanced TAU (psychoeducation booklet, monitoring and improved access to support) (6) TAU (7) Enhanced TAU (GP training) (8)-(9) TAU (10) Waitlist (11) TAU (12) Enhanced TAU (non-specific emotional support and mothercraft advice) (13) Enhanced TAU (home visits) (14) Enhanced TAU (single session post-delivery discussion)	(1)-(2) Listening visits	Relational constructivist therapy	Support group

Note. Abbreviations: BDI=Beck Depression Inventory; CIS-R=Computerised version of the Clinical Interview Schedule – Revised; DSM-III-R=Diagnostic and Statistical Manual of Mental Disorders, Third Edition, Revised; DSM-IV=Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition; EPDS=Edinburgh Postnatal Depression Scale; ICD-10=International Classification of Diseases, Tenth Revision; MCMI-III=Millon Clinical Multiaxial Inventory-III; MDD=Major depressive disorder; NR=Not reported; SCID=Structured Clinical Interview for DSM Disorders; TAU=Treatment as usual.
[1] Intensity: Low intensity (<8 sessions of contact with healthcare professional); Moderate intensity (8-15 sessions of contact with healthcare professional); High intensity (≥16 sessions of contact with healthcare professional).
[2] Time points: post-treatment or first measurement; Short-term follow-up (9-16 weeks post-intervention); Intermediate follow-up (17-24 weeks post-intervention); Long-term follow-up (25-103 weeks post-intervention follow-up); Very long-term follow-up (≥104 weeks).
[3] Four-armed trial: IPT; Listening visits; Mother-infant relationship intervention; TAU. Listening visits and Mother-infant relationship intervention comparisons extracted below. Demographic data is based on whole sample.
[4] Four-armed trial: CBT; Directive counselling (Individual); Directive counselling (Group); TAU. Directive counselling comparisons extracted below. Demographic data is based on whole sample.
[5] Three-armed trial that includes both prevention (whole sample) and treatment ('depressed' subgroup) data: CBT; Listening visits; TAU. Listening visits versus TAU comparison extracted below. Demographic data is based on all three arms.

Table 122: Study information table for trials included in the meta-analysis of self-help or facilitated self-help versus any alternative management strategy

	Self-help or faciltated self-help versus TAU	Post-miscarriage self-help versus TAU	Post-miscarriage facilitated self-help versus TAU
Total number of trials (number of participants)	3 (1,136)	2 (255)	1 (171)
Study ID	(1) OMAHEN2013A (2) OMAHEN2013C (3) MILGROM2011A	(1) KERSTING2011 (2) SWANSON2009[1]	SWANSON2009
Country	(1)-(2) UK (3) Australia	(1) Germany (2) US	US
Mean age of participants (years)	(1) 32.3 (2) NR (3) 32.3	(1) 34.3 (2) 32.4	32.4
Baseline diagnostic status	(1) Symptoms of depression (EPDS >12) (2) Diagnosis of MDD (diagnostic clinical assessment [on telephone] for ICD-10) (3) Subthreshold symptoms of depression (EPDS=8.9)	(1) Subthreshold symptoms of PTSD (IES=34) (2) Symptoms of depression (CES-D=21)	Symptoms of depression (CES-D=21)
Timing of intervention	(1)-(2) Postnatal (3) Antenatal	(1)-(2) Post-miscarriage	Post-miscarriage
Mode of delivery	(1) Internet delivery and online (chat room) support (2) Internet delivery and telephone support (3) Workbook delivery and telephone support	(1) Internet (2) Video and workbook	Video and workbook delivery and face-to-face support
Format	(1)-(3) Individual	(1)-(2) Individual	Individual
Intensity (number of sessions)	(1) Low (median support sessions=1-2 [11 internet sessions]) (2) Moderate (mean support sessions=8 [mean internet sessions=5]) (3) Moderate (support sessions=8 [workbook units=8])	(1) Low (no contact [10 written assignments]) (2) Low (no contact [3 workbook and video sessions])	Low (1 support session [3 workbook and video sessions])
Length of intervention (weeks)	(1) 15 (2) NR (3) 8	(1) 5 (2) 11	11
Time points	(1)-(3) Post-treatment	(1) Post-treatment (2) Post-treatment; Long follow-up	Post-treatment; Long follow-up
Setting	(1)-(2) Internet	(1) Internet	Home (for support)

	(3) Workbook	(2) Video and workbook	
Intervention	(1) (Facilitated) self-help (2)-(3) Facilitated self-help	(1)-(2) Self-help	Facilitated self-help
Comparison	(1)-(3) TAU	(1) Waitlist (2) TAU	TAU

Note. Abbreviations: CES-D=Center for Epidemiologic Studies Depression Scale; IES=Impact of Events Scale; NR=Not reported; TAU=Treatment as usual; EPDS=Edinburgh Postnatal Depression Scale; ICD-10=International Classification of Diseases, Tenth Revision; MDD=Major depressive disorder

[1] Four-armed trial: post-miscarriage self-help; Post-miscarriage facilitated self-help; Post-miscarriage counselling; TAU. Post-miscarriage counselling comparison extracted below. Demographic data is based on whole sample.

Table 123: Study information table for trials included in the meta-analysis of counselling versus any alternative management strategy

	Listening visits (non-directive counselling) versus TAU	Directive counselling versus TAU	Post-miscarriage counselling versus TAU/Enhanced TAU	Post-traumatic birth counselling versus TAU
Total number of trials (number of participants)	5 (1,018)	1 (146)	3 (269)	1 (103)
Study ID	(1) COOPER2003/MURRAY2003[1] (2) HOLDEN1989 (3) MORRELL2009A/2009B/2011/ BRUGHA2011[2] (4) WICKBERG1996 (5) WIGGINS2005	MILGROM2005[3]	(1) NEUGEBAUER2006 (2) NIKCEVIC2007 (3) SWANSON2009[4]	GAMBLE2005
Country	(1)-(3) UK (4) Sweden (5) UK	Australia	(1) US (2) UK (3) US	Australia
Mean age of participants (years)	(1) 27.7 (2) 26.2 (3) 30.9 (4) 28.4 (5) 29.6	29.7	(1) 29.7 (2) 35.3 (3) 32.4	28
Baseline diagnostic status	(1) Diagnosis of MDD (SCID for DSM-III-R) (2) Diagnosis of depression (Goldberg's standardised psychiatric interview for research diagnostic criteria) (3) Symptoms of depression (EPDS ≥12) (4) Diagnosis of MDD (interview by researcher and assessment with MADRS for DSM-III-R) (5) Subthreshold symptoms of depression (EPDS=8.9)	Diagnosis of minor depression or MDD (CIDI for DSM-IV)	(1) Symptoms of depression (100% HRSD>7. HRSD=16.5) (2) Symptoms of anxiety (HADS-A=8) (3) Symptoms of depression (CES-D=21)	Diagnosis of PTSD (MINI- PTSD for DSM-IV)
Timing of intervention	(1)-(5) Postnatal	Postnatal	(1)-(3) Post-miscarriage	Postnatal

Mode of delivery	(1)-(5) Face-to-face	Face-to-face	(1) Telephone (2)-(3) Face-to-face	Face-to-face
Format	(1)-(5) Individual	Individual or group	(1)-(3) Individual	Individual
Intensity (number of sessions)	(1)-(3) Moderate (8-10 sessions) (4) Low (6 sessions) (5) Moderate (10 sessions)	Moderate (11 sessions)	(1) Low (1-6 sessions) (2) Low (single session) (3) Low (3 sessions)	Low (2 sessions)
Length of intervention (weeks)	(1) 10 (2) 13 (3) 8 (4) 6 (5) 52	12	(1) 6 (2) Single sessions (3) 11	6
Time points	(1) Post-treatment; Intermediate follow-up; Long follow-up; Very long follow-up (2) Post-treatment (3) First measurement (4) Post-treatment (5) Post-treatment; Long follow-up	Post-treatment; Long follow-up	(1) Post-treatment (2) Post-treatment; Intermediate follow-up (3) Post-treatment; Long follow-up	Post-treatment
Setting	(1)-(5) Home	Clinic (primary)	(1) Telephone (2) Clinic (secondary) (3) Home	Face-to-face and telephone
Intervention	(1)-(2) Non-directive counselling (3) Listening visits (Person-centred approach) (4) Non-directive counselling (5) Listening visits	Directive counselling (individual and group counselling combined)	(1) Interpersonal counselling (2) Psychological counselling (+ medical investigations into causes of miscarriage) (3) Nurse-led counselling	Post-traumatic birth counselling
Comparison	(1)-(4) TAU (5) TAU (community support group and control group combined)	TAU	(1) TAU (2) Enhanced TAU (medical investigations into causes of miscarriage without counselling) (3) TAU	TAU
Note. Abbreviations: CES-D=Center for Epidemiologic Studies Depression Scale; DSM-III-R=Diagnostic and Statistical Manual of Mental Disorders, Third Edition, Revised; DSM-IV=Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition; EPDS=Edinburgh Postnatal Depression Scale; HADS-A=Hospital Anxiety and Depression Scale-Anxiety; HRSD=Hamilton Rating Scale for Depression; ICD-10=International Classification of Diseases, Tenth				

Revision; MADRS=Montgomery-Åsberg Depression Rating Scale; MDD=Major depressive disorder; MINI- PTSD=Mini-International Neuropsychiatric Interview-Post-Traumatic Stress Disorder; NR=Not reported; SCID=Structured Clinical Interview for DSM Disorders; TAU=Treatment as usual.
[1]Four-armed trial: IPT; Listening visits; Mother-infant relationship intervention; TAU. IPT comparison extracted above and Mother-infant relationship intervention comparison extracted below. Demographic data is based on whole sample
[2]Three-armed trial that includes both prevention (whole sample) and treatment ('depressed' subgroup) data: CBT; Listening visits; TAU. CBT versus TAU comparison extracted above. Demographic data is based on all three arms
[3]Four-armed trial: CBT; Directive counselling (Individual); Directive counselling (Group); TAU. CBT comparison extracted above. Demographic data is based on whole sample.
[4]Four-armed trial: post-miscarriage self-help; Post-miscarriage facilitated self-help; Post-miscarriage counselling; TAU. Post-miscarriage self-help and facilitated self-help comparisons extracted above. Demographic data is based on whole sample.

Table 124: Study information table for trials included in the meta-analysis of education or support versus any alternative management strategy

	Social support versus TAU	**Psychologically (CBT/IPT)-informed psychoeducation versus TAU or Enhanced TAU**	**IPT-informed psychoeducation versus non-mental health-focused education and support**	**Non-mental health-focused education and support versus TAU**	**Home visits versus TAU or Enhanced TAU**	**Pre-delivery discussion/ psychoeducation versus TAU**
Total number of trials (number of participants)	4 (867)	16 (2,955)	1 (38)	1 (331)	5 (1,616)	2 (547)
Study ID	(1) CHEN2000 (2) DENNIS2003 (3) DENNIS2009/2010 (4) LETOURNEAU2011	(1) AUSTIN2008 (2) BERNARD2011 (3) ELMOHANDES2008 (4) GAO2010/2012 (5) HAGAN2004 (6) HONEY2002 (7) KOZINSZKY2012 (8) LE2011	SPINELLI2003	KAAYA2013	(1) ARMSTRONG1999/ 2000/FRASER2000 (2) DUGGAN2007/ CALDERA2007 (3) DUGRAVIER2013/ GUEDENEY2013	(1) ROUHE2012/ SALMELAARO2012 (2) SAISTO2001

		(9) LEUNG2012 (10) MUNOZ2007/URIZAR2011 (11) SILVERSTEIN2011 (12) TANDON2011/2014/ MENDELSON2013 (13) TIMPANO2011 (14) ZLOTNICK2001 (15) ZLOTNICK2006 (16) ZLOTNICK2011			(4) ROMAN2009 (5) TAMAKI2008	
Country	(1) Taiwan (2)-(4) Canada	(1) Australia (2) US (3) US (4) China (5) Australia (6) UK (7) Hungary (8) US (9) China (10)-(16) US	US	Tanzania	(1) Australia (2) US (3) France (4) US (5) Japan	(1)-(2) Finland
Mean age of participants (years)	(1) 29.1 (2)-(4) NR	(1) 31.4 (2) 32.7 (3) 24.6 (4) 28.4 (5) Median: 29 (6) 27.9 (7) 27.3 (8) 25.4 (9) 31.2 (10) 24.9 (11) 27 (12) 23 (13) 27.3 (14) 23.4 (15) 22.4 (16) 23.8	28.7	26	(1) 26.2 (2) 23.6 (3) 22.3 (4) NR (5) 33.8	(1) 29.4 (2) 31.6

Baseline diagnostic status	(1) Symptoms of depression (BDI ≥10) (2)-(3) Symptoms of depression (EPDS >9) (4) Symptoms of depression (EPDS >12)	(1) Subthreshold symptoms of depression (EPDS=8) (2) Subthreshold symptoms of depression (BDI-II=13) (3) 51% of sample had symptoms of depression (Hopkins Symptom Checklist [HSCL]: Sum/20>0.75 depression) (4) Subthreshold symptoms of depression (EPDS=8) (5) Subthreshold symptoms of depression (median EPDS=8) (6) Symptoms of depression (EPDS>12) (7) Symptoms of depression (Leverton Questionnaire≥12) (8) Symptoms of depression (CES-D>16 and/or [family] history of depression) (9) Subthreshold symptoms of depression (EPDS=8) (10) Symptoms of depression (CES-D=16) (11) Symptoms of depression (QIDS=9) (12) Symptoms of depression (BDI=15) (13) Subthreshold symptoms of OCD (Obsessive Beliefs Questionnaire =170)	Diagnosis of MDD (SCID for DSM-IV)	73% of sample had symptoms of depression (HSCL-25>1.06)	(1) Subthreshold symptoms of depression (EPDS=8.7) (2) 57% of sample had symptoms of depression (CES-D >15) (3) Symptoms of depression (EPDS=11) (4) Symptoms of depression (CES-D=20) (5) Diagnosis of depression (SCID for DSM-IV)	(1) Symptoms of primary tokophobia (W-DEQ-A sum score=>100) (2) Symptoms of primary (51%) or secondary (49%) tokophobia (scored=>5/10 on study-specific fear of childbirth scale or request for caesarean)

		(14) 57% of sample had symptoms of depression (BDI>10; BDI=11) (15) Symptoms of depression (BDI=16) (16) Subthreshold symptoms of depression (EPDS=8)				
Timing of intervention	(1)-(4) Postnatal	(1) Antenatal (2) Postnatal (3)-(4) Antenatal and postnatal (5)-(6) Postnatal (7)-(8) Antenatal and postnatal (9) Antenatal (10) Antenatal and postnatal (11) Postnatal (12) Antenatal or postnatal (13)-(14) Antenatal (15)-(16) Antenatal and postnatal	Antenatal	Antenatal and postnatal	(1) Postnatal (2) Antenatal and postnatal or postnatal-only (3) Antenatal and postnatal (4) Antenatal and postnatal (5) Postnatal	(1) Antenatal and postnatal (2) Antenatal
Mode of delivery	(1) Face-to-face (2)-(3) Telephone (4) Face-to-face and telephone	(1)-(3) Face-to-face (4) Face-to-face and telephone (5)-(16) Face-to-face	Face-to-face	Face-to-face	(1)-(5) Face-to-face	(1)-(2) Face-to-face
Format	(1) Group (2)-(4) Individual	(1) Group (2)-(3) Individual (4) Individual and group (5)-(10) Group (11) Individual (12)-(15) Group (16) Individual	Group	Group	(1)-(5) Individual	(1) Group (2) Individual

Intensity (number of sessions)	(1) Low (4 sessions) (2) Low (no contact with professionals [5 sessions of peer support]) (3)-(4) Low (no contact with professionals [9 sessions of peer support])	(1)-(5) Low (3-6 sessions) (6) Moderate (8 sessions) (7)-(9) Low (4-6 sessions) (10) Moderate (8 sessions) (11)-(16) Low (3-6 sessions)	High (16 sessions)	Low (6 sessions)	(1) High (18 sessions) (2) High (42 sessions) (3) Low (7 sessions) (4) High (24 sessions) (5) Low (4 sessions)	(1)-(2) Low (6-7 sessions)
Length of intervention (weeks)	(1) 4 (2) 8 (3) NR (4) 12	(1) 6 (2) 3 (3)-(4) NR (5) 6 (6) 8 (7) 4 (8) 8 (9) 4 (10) 12 (11) 8 (12)-(13) 6 (14)-(16) 4	16	6	(1) 52 (2) 104 (3) 22 (4) NR (5) 5	(1) NR (2) 14
Time points	(1)-(2) Post-treatment (3) Post-treatment; Short follow-up (4) Post-treatment	(1) First measurement; Intermediate follow-up (2)-(3) Post-treatment (4) Post-treatment; Short follow-up (5) Post-treatment; Intermediate follow-up; Long follow-up (6) Post-treatment; Long follow-up (7) First measurement (8) Post-treatment; Intermediate follow-up; Long follow-up	Post-treatment	Post-treatment	(1) Post-treatment; First measurement (2) Post-treatment (3) Post-treatment; First measurement (4)-(5) Post-treatment	(1)-(2) Mid-treatment; Post-treatment; First measurement

		(9) Post-treatment; Intermediate follow-up (10) Post-treatment; Short follow-up; Intermediate follow-up; Long follow-up (11) First measurement; Intermediate follow-up (12)-(13) Post-treatment; Short follow-up; Long follow-up (14) Post-treatment (15)-(16) First measurement				
Setting	(1) NR (2)-(3) Telephone (4) Home and telephone	(1)-(3) NR (4) Clinic (primary) and telephone (5)-(10) NR (11) Hospital or home (12)-(16) NR	NR	Hospital	(1)-(5) Home	(1) NR (2) Hospital
Intervention	(1) Support group (2)-(3) Peer-mediated support (4) Peer-mediated support (with mother–infant relationship intervention content)	(1)-(3) CBT-informed psychoeducation (4) IPT-informed psychoeducation (5)-(6) CBT-informed psychoeducation (7) CBT- and IPT-informed psychoeducation (8) CBT-informed psychoeducation (9) IPT-informed psychoeduation (10)-(13) CBT-informed psychoeducation (14)-(16) IPT-informed psychoeduation	IPT-informed psychoeducation	Non-mental health-focused education and support (group counselling intervention for HIV-positive women)	(1)-(5) Home visits	(1) CBT-informed psychoeducation (2) Pre-delivery discussion/IPT-informed psychoeducation
Comparison	(1)-(3) TAU (4) Waitlist	(1) Enhanced TAU (psychoeducation booklet)	Non-mental health-focused	TAU	(1)-(3) TAU	(1)-(2) TAU

		(2)-(3) TAU (4) Enhanced TAU (non-mental health-focused education and support group) (5)-(6) TAU (7) Enhanced TAU (non-mental health-focused education and support group) (8)-(11) TAU (12) Enhanced TAU (psychoeducation booklet) (13) Enhanced TAU (psychoeducation group [without CBT component]) (14)-(16) TAU	education and support (group)		(4) Enhanced TAU (Medicaid enhanced prenatal/postnatal services) (5) TAU	

Note. Abbreviations: BDI=Beck Depression Inventory; CES-D=Center for Epidemiologic Studies Depression Scale; DSM-IV=Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition; EPDS=Edinburgh Postnatal Depression Scale; HSCL=Hopkins Symptom Checklist; =MDD=Major depressive disorder; NR=Not reported; =QIDS=Quick Inventory of Depressive Symptoms; SCID=Structured Clinical Interview for DSM Disorders; TAU=Treatment as usual; W-DEQ-A=Wijma Delivery Expectancy Questionnaire

Table 125: Study information table for trials included in the meta-analysis of mother–infant relationship interventions versus any alternative management strategy

	Mother-infant relationship interventions versus TAU or Enhanced TAU	Mother-infant relationship intervention (video feedback) versus mother–infant relationship intervention (verbal feedback)	Mother-infant relationship intervention (+ facilitated self-help for eating disorders) versus listening visits (+ facilitated self-help for eating disorders)	Co-parenting intervention versus Enhanced TAU
Total number of trials (number of participants)	6 (691)	1 (51)	1 (80)	1 (29)
Study ID	(1) COOPER2003/ MURRAY2003[1] (2) HOROWITZ2001 (3) SALOMONSSON2011 (4) SLEED2013 (5) VANDOESUM2008/ KERSTENALVAREZ2010 (6) ZELKOWITZ2008/2011/ FEELEY2012	BILSZTA2012[2]	STEIN2006	MISRI2000
Country	(1) UK (2) US (3) Sweden (4) UK (5) Netherlands (6) Canada	Australia	UK	Canada
Mean age of participants (years)	(1) 27.7 (2) 31 (3) 33.6 (4) 26.8 (5) 30	NR	Median=30	33.2

	(6) 30.9			
Baseline diagnostic status	(1) Diagnosis of MDD (SCID for DSM-III-R) (2) Symptoms of depression (EPDS=>10) (3) Symptoms of depression (EPDS=12) (4) Subthreshold symptoms of depression (CES-D=15) (5) 95% of sample had diagnosis of a major depressive episode or dysthymia (MINI for DSM-IV) (6) Symptoms of depression (EPDS=14), anxiety (STAI=47), and/or PTSD (PPQ=6)	Diagnosis of MDD (DSM-IV [assessment tool not specified])	Diagnosis of eating disorder (psychiatric interview for DSM-IV)	Diagnosis of MDD (MINI for DSM-IV)
Timing of intervention	(1)-(6) Postnatal	Postnatal	Postnatal	Postnatal
Mode of delivery	(1)-(6) Face-to-face	Face-to-face	Face-to-face	Face-to-face
Format	(1)-(3) Individual (4) Group (5)-(6) Individual	Individual	Individual	Individual
Intensity (number of sessions)	(1) Moderate (10 sessions) (2) Low (3 sessions) (3) High (29 sessions) (4) Low (7 sessions) (5) Moderate (8-10 sessions) (6) Low (6 sessions)	Low (3 sessions)	Moderate (12 sessions)	Low (4 sessions)
Length of intervention (weeks)	(1) 10 (2) 18 (3) 12 (4) 4 (5) 15 (6) NR	3	30	6

Time points	(1) Post-treatment; Intermediate follow-up; Long follow-up; Very long follow-up (2) Post-treatment (3) First measurement (4) Post-treatment (5) Post-treatment; Long follow-up; Very long follow-up (6) Post-treatment; First measurement; Intermediate follow-up	Post-treatment	Post-treatment	Post-treatment
Setting	(1)-(2) Home (3) Clinic (secondary) (4) Prison (5) Home (6) NR	Hospital	Home	Clinic (primary)
Intervention	(1)-(2) Mother-infant relationship intervention (3) Mother-infant psychotherapy (4)-(6) Mother-infant relationship intervention	Mother-infant relationship intervention (with video feedback)	Mother-infant relationship intervention (and facilitated self-help aimed at the eating disorder)	Co-parenting intervention
Comparison	(1) TAU (2) Enhanced TAU (video assessment without coaching) (3)-(4) TAU (5) Enhanced TAU (telephone support) (6) Enhanced TAU (non-mental health-focused education and support [booklet about infant care])	Mother-infant relationship intervention (with verbal feedback)	Listening visits (and facilitated self-help aimed at the eating disorder)	Enhanced TAU (monitoring)

Note. Abbreviations: CES-D=Center for Epidemiologic Studies Depression Scale; DSM-III-R=Diagnostic and Statistical Manual of Mental Disorders, Third Edition, Revised; DSM-IV=Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition; =EPDS=Edinburgh Postnatal Depression Scale; MDD=Major depressive disorder; MINI=Mini-International Neuropsychiatric Interview; NR=Not reported; PPQ=Perinatal PTSD Questionnaire; SCID=Structured Clinical Interview for DSM Disorders; STAI=State-Trait Anxiety Inventory; TAU=Treatment as usual

[1]Four-armed trial: IPT; Listening visits; Mother-infant relationship intervention; TAU. IPT and Listening visits comparisons extracted above. Demographic data is based on whole sample
[2]This was a three-armed trial which also included a TAU arm, however, data could not be extracted for the TAU arm due to non-random assignment to that condition

Table 126: Study information table for trials included in the meta-analysis of other psychosocial interventions versus any alternative management strategy

	Infant sleep training (controlled crying) versus TAU or Enhanced TAU	Music therapy during birth versus TAU	Psychosomatic interventions versus TAU	Mindfulness training versus TAU or Enhanced TAU
Total number of trials (number of participants)	2 (394)[1]	1 (161)	2 (276)	2 (81)
Study ID	(1) HISCOCK2002 (2) HISCOCK2007/2008	SIMAVLI2014	(1) ORTIZCOLLADO2014 (2) WEIDNER2010	(1) GUARDINO2014 (2) VIETEN2008
Country	(1)-(2) Australia	Turkey	(1) Spain and France (2) Germany	(1)-(2) US
Mean age of participants (years)	(1)-(2) NR	23.8	(1) 29.3 (2) 28	(1) 33.1 (2) 33.9
Baseline diagnostic status	(1) Symptoms of depression (EPDS=>10) (2) HISCOCK2007: Symptoms of depression (EPDS >9). HISCOCK2008: Subthreshold symptoms of depression (EPDS=8)	Subthreshold symptoms of depression (EPDS=8)	(1) Symptoms of depression (EPDS=11) (2) Symptoms of anxiety (HADS-A=9)	(1) Symptoms of Anxiety (STAI-State=45) (2) Symptoms of depression (31% of sample CES-D>16. CES-=16.8)
Timing of intervention	(1)-(2) Postnatal	During delivery	(1)-(2) Antenatal	(1)-(2) Antenatal
Mode of delivery	(1)-(2) Face-to-face	CD	(1)-(2) Face-to-face	(1)-(2) Face-to-face
Format	(1)-(2) Individual	Individual	(1) Group (2) Individual	(1)-(2) Group
Intensity (number of sessions)	(1)-(2) Low (2-3 sessions)	Low (1 session)	(1) Moderate (10 sessions) (2) Low (1-5 sessions)	(1)-(2) Low (5-7 sessions)

Length of intervention (weeks)	(1) 6 (2) 2	Single session	(1) 10 (2) NR	(1) 6 (2) 8
Time points	(1) Post-treatment; Short follow-up (2) Post-treatment; First measurement; Short follow-up; Long follow-up	Post-treatment	(1)-(2) First measurement	(1)-(2) Post-treatment
Setting	(1)-(2) Clinic (primary)	Hospital	(1)-(2) Hospital	(1) Clinic (secondary) (2) Hospital
Intervention	(1)-(2) Controlled crying (or camping out)	Music therapy during birth	(1)-(2) Psychosomatic intervention	(1)-(2) Mindfulness training
Comparison	(1) Enhanced TAU (non-mental health-focused education and support [booklet about infant sleep]) (2) TAU	TAU	(1)-(2) TAU	(1) Enhanced TAU (non-mental health-focused education and support [book]) (2) Waitlist

Note. Abbreviations: EPDS=Edinburgh Postnatal Depression Scale; HADS-A=Hospital Anxiety and Depression Scale-Anxiety; NR=Not reported; STAI=State-Trait Anxiety Inventory; TAU=Treatment as usual

[1]Where possible data is only extracted for the 'depressed' subgroup (EPDS>9/10), however, this is not possible for HISCOCK2008 so for this paper whole sample data is extracted

Table 127: Study information table for trials included in the meta-analysis of combined psychosocial and physical interventions

	Combined social support and physical exercise versus Enhanced TAU	**Social support versus physical excercise**
Total number of trials (number of participants)	1 (20)	1 (24)
Study ID	ARMSTRONG2003	ARMSTRONG2004
Country	Australia	Australia
Mean age of participants (years)	NR	NR
Baseline diagnostic status	100% of sample had symptoms of depression (EPDS=>12)	100% of sample had symptoms of depression (EPDS=>12)
Timing of intervention	Postnatal	Postnatal
Mode of delivery	Face-to-face	Face-to-face
Format	Group	Group
Intensity (number of sessions)	High (48 sessions)	Moderate (12 sessions)
Length of intervention (weeks)	12	12
Time points	Post-treatment	Post-treatment
Setting	Community	Community
Intervention	Pram walking with informal gathering	Social support group
Comparison	Telephone support (at midpoint)	Pram walking exercise programme
Note. Abbreviations: NR=Not reported; TAU=Treatment as usual		

Table 128: Study information table for systematic reviews and primary RCTs included in the review of psychosocial interventions for alcohol and substance misuse

Cochrane review	Primary objective	Inclusion criteria	Included studies	Additional studies
STADE2009B	Determine the effectiveness of either psychological or educational interventions, or both, for reducing prenatal consumption of alcohol among pregnant women, or women planning for pregnancy.	Pregnant women/women planning pregnancy who consume alcohol, and who are participating in studies examining psychological or educational interventions to reduce alcohol	• Chang and colleagues (1999, 2000) • Handmaker and colleagues (1999) • O'Connor and Whaley (2007) • Reynolds and colleagues (1995) Awaiting assessment: • Chang and colleagues (2005, 2006)	MARAIS2011 OSTERMAN2012 OSTERMAN2014

TERPLAN2007	Evaluate the effectiveness of psychosocial interventions in pregnant women enrolled in illicit drug treatment programmes on birth and neonatal outcomes, on attendance and retention in treatment, as well as on maternal and neonatal drug abstinence.	Pregnant women enrolled in illicit drug treatment programs (illegal substances such as cannabis, heroin, cocaine, amphetamines) Women on methadone are also included	• Carrol and colleagues (1995) • Elk and colleagues (1998) • Haug and colleagues (2004) • Jones and colleagues (2000) • Jones and colleagues (2001) • Mullins and colleagues (2004) • O'Neill and colleagues (1996) • Silverman and colleagues (2001) • Svikis and colleagues (1997)	WINHUSEN2008 YONKERS2012
TURNBULL2012	Determine the effectiveness of home visits on improving outcome for pregnant or postpartum women with a drug or alcohol problem	Pregnant or postpartum women with an alcohol or drug problem.	• Bartu and colleagues (2006) • Black and colleagues (1994) • Butz and colleagues (1998, 2001) • Dakof and colleagues (2003) • Grant and colleagues (1996a, 1996b, 2005)/Ernst and colleagues (1999)/Kartin	None

			and colleagues (2002) • Quinlivan and colleagues (2003) • Schuler and colleagues (2000, 2002a, 2002b, 2003)/Ackerman and colleagues (2008)/Kettinger and colleagues (2000)/Nair and colleagues (2002, 2003, 2008)	
No relevant Cochrane review	Determine the effectiveness of psychologically-informed psychoeducation for improving outcomes for women who show at-risk drinking in the postnatal period	Women in the postnatal period who tested positive for at-risk drinking	Not applicable	FLEMING2008
No relevant Cochrane review	Determine the effectiveness of self-help on reducing illicit drug use for women in the postnatal period	Women in the postnatal period who met criteria for illicit drug use in the month before becoming pregnant	Not applicable	ONDERSMA2014
No relevant cochrane review	Determine the long-term efficacy of contingency management on continued illicit drug abstinence in the postnatal period	Long-term follow-up of pregnant women enrolled in illicit drug treatment program (heroin, cocaine, methodone maintenance treatment)	Not applicable	SILVERMAN2002

7.5.4 Clinical evidence for effects on depression outcomes (by intervention)

Summary of findings can be found in the tables presented in this section. The full GRADE evidence profiles and associated forest plots can be found in Appendix 22 and Appendix 19, respectively.

Depression: structured psychological interventions (CBT or IPT) versus treatment as usual or enhanced treatment as usual

Very low to high quality evidence from up to ten studies (N=1,508) showed that structured psychological interventions (CBT or IPT) were more effective than treatment as usual or enhanced treatment as usual (using both ITT and available case analysis) in reducing depression diagnosis ($p < 0.0001$), depression symptomatology ($p \leq 0.0004$) and depression mean scores ($p < 0.00001$) at post-treatment, with large to moderate effects observed for all outcomes and some low quality evidence for maintained moderate to large effects at short-term follow-up (9-16 weeks post-intervention; $p < 0.01$) (Table 129). At intermediate follow-up periods (17-24 weeks post-intervention) there was evidence for moderate benefits associated with structured psychological interventions, however, confidence that these were true measures of effect was low to very low due to wide confidence intervals including the possibility of both no effect and clinically significant benefits for depression diagnosis (available case analysis) and depression mean scores (p=0.08-0.41) and in the case of the ITT analysis of depression diagnosis the 95% CI spans the thresholds for harm, no effect and benefit (p=0.23). At longer-term follow-ups (>24 weeks post-intervention), the evidence for structured psychological interventions is very inconsistent with point estimates of effect in favour of CBT or IPT for depression symptomatology (p=0.41-0.59), but in favour of treatment as usual or enhanced treatment as usual for depression diagnosis (p=0.02-0.25) (Table 129).

Table 129: Summary of findings table for effects of structured psychological interventions (CBT or IPT) compared with treatment as usual or enhanced treatment as usual on depression outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk	Corresponding risk	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Control	Depression: Structured psychological interventions (CBT or IPT) versus TAU/Enhanced TAU				
	Study population					

Depression diagnosis post-treatment – ITT analysis SCID or CIS-R Follow-up: 12-44 weeks	652 per 1000	313 per 1000 (254 to 391)	RR 0.48 (0.39 to 0.6)	1307 (6 studies)	⊕⊕⊕⊕ high	
	Moderate					
	687 per 1000	330 per 1000 (268 to 412)				
Depression diagnosis post-treatment – available case analysis SCID or CIS-R Follow-up: 12-44 weeks	Study population		RR 0.38 (0.24 to 0.58)	1,066 (5 studies)	⊕⊕⊖⊖ low[1]	
	602 per 1000	229 per 1000 (145 to 349)				
	Moderate					
	615 per 1000	234 per 1000 (148 to 357)				
Depression symptomatology Post-treatment – ITT analysis EPDS ≥10/EPDS ≥12/Treatment non-response (baseline-endpoint decrease <4 points and EPDS >13)/Treatment non-response (<50% improvement) or BDI ≥16 or BDI-II ≥14 Follow-up: 6-44 weeks	Study population		RR 0.69 (0.56 to 0.85)	969 (10 studies)	⊕⊕⊖⊖ low[2,3]	
	643 per 1000	444 per 1000 (360 to 547)				
	Moderate					
	626 per 1000	432 per 1000 (351 to 532)				
Depression symptomatology Post-treatment – available case analysis EPDS ≥10/EPDS ≥12/Treatment non-response (baseline-endpoint decrease <4 points and EPDS >13) or BDI ≥16 or BDI-II ≥14 Follow-up: 6-16 weeks	Study population		RR 0.62 (0.53 to 0.73)	702 (9 studies)	⊕⊕⊕⊕ high	
	559 per 1000	347 per 1000 (296 to 408)				
	Moderate					
	588 per 1000	365 per 1000 (312 to 429)				
Depression mean scores post-treatment – ITT analysis EPDS or BDI-II Follow-up: 6-44 weeks		The mean depression mean scores post-treatment – ITT analysis in the intervention groups was 1.31 standard deviations lower (2.36 to 0.26 lower)		306 (5 studies)	⊕⊖⊖⊖ very low[1,4]	SMD -1.31 (-2.36 to -0.26)
Depression mean scores post-treatment – available case analysis EPDS, BDI, BDI-II or HRSD Follow-up: 6-16 weeks		The mean depression mean scores post-treatment – available case analysis in the intervention groups was 0.6 standard		1,508 (10 studies)	⊕⊕⊕⊖ moderate[2]	SMD -0.6 (-0.8 to -0.4)

		deviations lower (0.8 to 0.4 lower)				
Depression diagnosis Short Follow-up (9-16 weeks post-intervention) – ITT analysis SCID Follow-up: mean 28 weeks	Study population 435 per 1000 Moderate 435 per 1000	170 per 1000 (83 to 348) 170 per 1000 (83 to 348)	RR 0.39 (0.19 to 0.8)	93 (1 study)	⊕⊕⊖⊖ low[5]	
Depression symptomatology Short Follow-up (9-16 weeks post-intervention) – ITT analysis BDI-II ≥14 Follow-up: mean 29 weeks	Study population 560 per 1000 Moderate 560 per 1000	498 per 1000 (302 to 823) 498 per 1000 (302 to 823)	RR 0.89 (0.54 to 1.47)	55 (1 study)	⊕⊕⊖⊖ low[5,6]	
Depression symptomatology Short Follow-up (9-16 weeks post-intervention) – available case analysis BDI-II ≥14 Follow-up: mean 29 weeks	Study population 667 per 1000 Moderate 667 per 1000	380 per 1000 (207 to 713) 380 per 1000 (207 to 714)	RR 0.57 (0.31 to 1.07)	42 (1 study)	⊕⊕⊖⊖ low[5]	
Depression mean scores Short Follow-up (9-16 weeks post-intervention) – ITT analysis EPDS or BDI-II Follow-up: 28-29 weeks		The mean depression mean scores short follow-up (9-16 weeks post-intervention) – ITT analysis in the intervention groups was 1.84 standard deviations lower (4.31 lower to 0.64 higher)		148 (2 studies)	⊕⊖⊖⊖ very low[1,4,6]	SMD -1.84 (-4.31 to 0.64)
Depression mean scores Short Follow-up (9-16 weeks post-intervention) – available case analysis EPDS or BDI-II Follow-up: 21-29 weeks		The mean depression mean scores short follow-up (9-16 weeks post-intervention) – available case analysis in the intervention groups was 0.66 standard deviations lower (1.14 to 0.18 lower)		89 (2 studies)	⊕⊕⊖⊖ low[4]	SMD -0.66 (-1.14 to -0.18)
Depression diagnosis intermediate follow-up (17-24 weeks post-intervention) – ITT analysis CIS-R or SCID	Study population 471 per 1000 Moderate 572 per 1000	278 per 1000 (113 to 665) 337 per 1000 (137 to 807)	RR 0.59 (0.24 to 1.41)	138 (2 studies)	⊕⊖⊖⊖ very low[5,6,7]	

Follow-up: mean 33 weeks						
Depression diagnosis intermediate follow-up (17-24 weeks post-intervention) – available case analysis CIS-R or SCID Follow-up: mean 33 weeks	Study population		RR 0.5 (0.23 to 1.08)	118 (2 studies)	⊕⊕⊖⊖ low[5,6]	
	373 per 1000	186 per 1000 (86 to 403)				
	Moderate					
	474 per 1000	237 per 1000 (109 to 512)				
Depression mean depression scores Intermediate follow-up (17-24 weeks post-intervention) – available case analysis EPDS Follow-up: mean 33 weeks		The mean depression mean depression scores intermediate follow-up (17-24 weeks post-intervention) – available case analysis in the intervention groups was 0.51 standard deviations lower (1.72 lower to 0.7 higher)		118 (2 studies)	⊕⊖⊖⊖ very low[1,4,6]	SMD -0.51 (-1.72 to 0.7)
Depression diagnosis Long Follow-up (25-103 weeks post-intervention) – ITT analysis SCID Follow-up: mean 78 weeks	Study population		RR 1.68 (0.95 to 2.98)	102 (1 study)	⊕⊕⊖⊖ low[5,6]	
	250 per 1000	420 per 1000 (237 to 745)				
	Moderate					
	250 per 1000	420 per 1000 (237 to 745)				
Depression diagnosis Long Follow-up (25-103 weeks post-intervention) – available case analysis SCID Follow-up: mean 78 weeks	Study population		RR 1.56 (0.73 to 3.33)	89 (1 study)	⊕⊕⊖⊖ low[5,6]	
	188 per 1000	292 per 1000 (137 to 624)				
	Moderate					
	188 per 1000	293 per 1000 (137 to 626)				
Depression symptomatology Long Follow-up (25-103 weeks post-intervention) – ITT analysis EPDS ≥10 Follow-up: mean 32 weeks	Study population		RR 0.71 (0.2 to 2.53)	37 (1 study)	⊕⊖⊖⊖ very low[5,6,8]	
	250 per 1000	178 per 1000 (50 to 632)				
	Moderate					
	250 per 1000	178 per 1000 (50 to 632)				
Depression symptomatology Long Follow-up (25-103 weeks post-	Study population		RR 0.4 (0.05 to 3.46)	33 (1 study)	⊕⊖⊖⊖ very low[5,6,8]	
	167 per 1000	67 per 1000 (8 to 577)				
	Moderate					

intervention) – available case analysis EPDS ≥10 Follow-up: mean 32 weeks	167 per 1000	67 per 1000 (8 to 578)				
Depression mean scores Long Follow-up (25-103 weeks post-intervention) – available case analysis EPDS or BDI Follow-up: 32-78 weeks		The mean depression mean scores long follow-up (25-103 weeks pot-intervention) – available case analysis in the intervention groups was 0.28 standard deviations lower (0.8 lower to 0.23 higher)		142 (3 studies)	⨁⨁⊖⊖ low[4,6]	SMD -0.28 (-0.8 to 0.23)
Depression diagnosis Very long Follow-up (>104 weeks post-intervention) – ITT analysis SCID Follow-up: mean 260 weeks	Study population 250 per 1000	480 per 1000 (278 to 832)	RR 1.92 (1.11 to 3.33)	102 (1 study)	⨁⨁⊖⊖ low[5]	
	Moderate 250 per 1000	480 per 1000 (278 to 832)				
Depression diagnosis Very long Follow-up (>104 weeks post-intervention) – available case analysis SCID Follow-up: mean 260 weeks	Study population 243 per 1000	212 per 1000 (90 to 506)	RR 0.87 (0.37 to 2.08)	70 (1 study)	⨁⨁⊖⊖ low[5,6]	
	Moderate 243 per 1000	211 per 1000 (90 to 505)				
Depression mean depression scores Very long Follow-up (>104 weeks post-intervention) – available case analysis EPDS Follow-up: mean 260 weeks		The mean depression mean depression scores very long follow-up (>104 weeks post-intervention) – available case analysis in the intervention groups was 0.17 standard deviations lower (0.67 lower to 0.33 higher)		62 (1 study)	⨁⨁⊖⊖ low[4,6]	SMD -0.17 (-0.67 to 0.33)
Negative thoughts/mood mean scores – available case analysis Automatic Thought Questionnaire Follow-up: mean 4 weeks		The mean negative thoughts/mood mean scores – available case analysis in the intervention groups was 0.94 standard		22 (1 study)	⨁⊖⊖⊖ very low[4,8]	SMD -0.94 (-1.83 to -0.04)

	deviations lower (1.83 to 0.04 lower)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] There was evidence of considerable heterogeneity between effect sizes
[2] There was evidence of moderate to substantial heterogeneity between effect sizes
[3] Papers omit data
[4] Total population size is less than 400 (a threshold rule-of-thumb)
[5] Total number of events is less than 300 (a threshold rule-of-thumb)
[6] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
[7] There was evidence of substantial heterogeneity between effect sizes
[8] Risk of bias due to statistically significant group differences at baseline

Depression: structured psychological interventions (CBT or IPT) versus alternative active intervention

There was no evidence for benefits associated with CBT relative to listening visits on mean depression symptoms at endpoint or first measurement (p=0.69; Table 130).

Table 130: Summary of findings table for effects of CBT compared with listening visits on depression outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	 Corresponding risk Depression: CBT versus listening visits	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Depression mean scores post-treatment – available case analysis BDI or EPDS Follow-up: mean 26 weeks		The mean depression mean scores post-treatment – available case analysis in the intervention groups was 0.06 standard deviations lower (0.33 lower to 0.22 higher)		301 (2 studies)	⊕⊕⊖⊖ low[1]	SMD -0.06 (-0.33 to 0.22)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Papers omit data

There was very low quality, single study (N=60) evidence for moderate benefits (p=0.04) associated with relational constructivist therapy over CBT on mean depression symptoms (Table 131).

Table 131: Summary of findings table for effects of CBT compared with relational constructivist therapy on depression outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	 Corresponding risk Depression: CBT versus relational constructivist therapy	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Depression mean scores post-treatment – available case analysis BDI		The mean depression mean scores post-treatment – available case analysis in the intervention groups was 0.53 standard deviations higher (0.01 to 1.05 higher)		60 (1 study)	⊕⊖⊖⊖ very low[1,2]	SMD 0.53 (0.01 to 1.05)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total population size is less than 400 (a threshold rule-of-thumb)
2 Paper omits data

There was no evidence for clinically or statistically significant effects of IPT relative to a support group on mean depression symptoms (p=0.11; Table 132).

Table 132: Summary of findings table for effects of IPT compared with support group on depression outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	 Corresponding risk Depression: IPT versus support group	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Depression mean scores post-treatment – available case analysis CES-D Follow-up: mean 12 weeks		The mean depression mean scores post-treatment – available case analysis in the intervention groups was 0.49 standard deviations lower (1.09 lower to 0.11 higher)		44 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	SMD -0.49 (-1.09 to 0.11)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Risk of bias due to statistically significant group differences at baseline
[2] Total population size is less than 400 (a threshold rule-of-thumb)
[3] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Depression: facilitated self-help versus treatment as usual

There was very low to high quality data from up to three studies (N=1,136) for moderate benefits (p <0.00001 to p=0.04) of facilitated self-help relative to treatment as usual for depression symptomatology (ITT and available case analysis) and mean depression symptoms (Table 133).

Table 133: Summary of findings table for effects of facilitated self-help compared with treatment as usual on depression outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Depression: Facilitated self-help versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Depression symptomatology Post-treatment – ITT Analysis BDI-II≥14 or EPDS>12 Follow-up: 15-20 weeks	Study population 817 per 1000	596 per 1000 (433 to 809)	RR 0.73 (0.53 to 0.99)	1,136 (3 studies)	⊕⊖⊖⊖ very low[1,2]	
	Moderate 762 per 1000	556 per 1000 (404 to 754)				
Depression symptomatology Post-treatment – available case analysis Beck Depression Inventory-II (BDI-II)≥14 or EPDS>12 Follow-up: 15-20 weeks	Study population 567 per 1000	329 per 1000 (250 to 437)	RR 0.58 (0.44 to 0.77)	503 (3 studies)	⊕⊕⊖⊖ low[2,3]	
	Moderate 586 per 1000	340 per 1000 (258 to 451)				
Depression mean scores post-treatment – available case analysis EPDS Follow-up: 15-17 weeks		The mean depression mean scores post-treatment – available case analysis in the intervention groups was 0.56 standard deviations lower (0.76 to 0.37 lower)		414 (2 studies)	⊕⊕⊕⊕ high	SMD -0.56 (-0.76 to -0.37)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] There was evidence of considerable heterogeneity between effect sizes
[2] Papers omit data
[3] Total number of events is less than 300 (a threshold rule-of-thumb)

Depression: post-miscarriage self-help or facilitated self-help versus treatment as usual

There was low quality, single study (N=78) evidence that post-miscarriage self-help was more effective than treatment as usual for depression symptomatology (analysed according to ITT [p=0.02] or available case [p=0.005] approaches) with moderate to large effects observed. However, the measure for depression symptomatology was treatment non-response (based on reverse scale rating of reliable change index) on the BSI Depression subscale rather than a depression-specific validated checklist. In addition, there was some discrepancy between dichotomous and continuous measures of depression. There was no evidence for clinically or statistically significant benefits (p=0.32-0.51) of post-miscarriage self-help or facilitated self-help on mean depression symptoms (
Table 134 and Table 135).

Table 134: Summary of findings table for effects of post-miscarriage self-help compared with treatment as usual on depression outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Depression: post-miscarriage self-help versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Depression symptomatology Post-treatment – ITT analysis BSI: Depression (Treatment non-response: reliable change index) Follow-up: mean 5 weeks	Study population		RR 0.65 (0.45 to 0.92)	78 (1 study)	⊕⊕⊖⊖ low1	
	758 per 1000	492 per 1000 (341 to 697)				
	Moderate					
	758 per 1000	493 per 1000 (341 to 697)				
Depression symptomatology Post-treatment – available case analysis BSI: Depression (Treatment non-response: reliable change index)	Study population		RR 0.44 (0.25 to 0.78)	59 (1 study)	⊕⊕⊖⊖ low1	
	692 per 1000	305 per 1000 (173 to 540)				
	Moderate					
	692 per 1000	304 per 1000 (173 to 540)				

Follow-up: mean 5 weeks				
Depression mean scores post-treatment – ITT analysis BSI: Depression or CES-D Follow-up: 5-12 weeks	The mean depression mean scores post-treatment – ITT analysis in the intervention groups was 0.3 standard deviations lower (1.19 lower to 0.6 higher)	250 (2 studies)	⊕⊖⊖⊖ very low2,3	SMD -0.3 (-1.19 to 0.6)
Depression mean scores Long Follow-up (25-103 weeks post-intervention) – ITT analysis CES-D Follow-up: mean 46 weeks	The mean depression mean scores long follow-up (25-103 weeks post-intervention) – ITT analysis in the intervention groups was 0.15 standard deviations lower (0.45 lower to 0.15 higher)	172 (1 study)	⊕⊕⊖⊖ low3	SMD -0.15 (-0.45 to 0.15)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total number of events is less than 300 (a threshold rule-of-thumb)
2 There was evidence of considerable heterogeneity between effect sizes
3 Total population size is less than 400 (a threshold rule-of-thumb)

Table 135: Summary of findings table for effects of post-miscarriage facilitated self-help compared with treatment as usual on depression outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Depression: post-miscarriage facilitated self-help versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Depression mean scores post-treatment – ITT analysis CES-D Follow-up: mean 12 weeks		The mean depression mean scores post-treatment – ITT analysis in the intervention groups was 0.13 standard deviations higher (0.17 lower to 0.43 higher)		171 (1 study)	⊕⊕⊖⊖ low[1]	SMD 0.13 (-0.17 to 0.43)
Depression mean scores Long Follow-up (25-103 weeks post-intervention) – ITT analysis CES-D Follow-up: mean 46 weeks		The mean depression mean scores long follow-up (25-103 weeks pot-intervention) – ITT analysis in the intervention groups was 0.1 standard deviations lower (0.4 lower to 0.2 higher)		171 (1 study)	⊕⊕⊖⊖ low[1]	SMD -0.1 (-0.4 to 0.2)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total population size is less than 400 (a threshold rule-of-thumb)

Depression: listening visits versus treatment as usual

When an available case method of analysis was adopted there was very low quality evidence from three studies (N=179) for moderate benefits (p=0.03) of listening visits on depression diagnosis (Table 136). However, there was no evidence for statistically

significant benefits of listening visits for depression diagnosis using an ITT data analysis approach (p=0.12) or for statistically or clinically significant effects of listening visits on depression symptomatology using an ITT or available case analysis approach (p=0.07-0.50), or for clinically significant effects on mean depression symptoms (p=0.001). In addition, at intermediate follow-up periods (17-24 weeks post-intervention) there was no evidence for statistically or clinically significant benefits on depression diagnosis using either data analysis method (p=0.62-0.91) or on depression mean symptoms (p=0.73). Moreover, at longer-term follow-ups the evidence for treatment effects is very inconsistent with no evidence for clinically or statistically significant benefits or harms of listening visits compared with treatment as usual on depression diagnosis at >104 week follow-up using an available case analysis (p=0.76) or depression symptomatology at 25-103 week follow-up (p=0.65-0.77) or mean depression symptoms at 25-103 week or >104 week follow-ups (p=0.45-0.49), but with point estimates suggestive of clinically significant harms(effects in favour of treatment as usual) on depression diagnosis at 25-103 week follow-up (p=0.18-0.26) and at >104 week follow-up (p=0.03).

Table 136: Summary of findings table for effects of listening visits compared with treatment as usual on depression outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Depression: Listening visits versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Depression diagnosis post-treatment – ITT analysis SCID Follow-up: mean 20 weeks	Study population 615 per 1000	455 per 1000 (314 to 665)	RR 0.74 (0.51 to 1.08)	100 (1 study)	⊕⊕⊖⊖ low[1,2]	
	Moderate 615 per 1000	455 per 1000 (314 to 664)				
Depression diagnosis post-treatment – available case analysis SCID or Goldberg's standardised psychiatric interview: Research diagnostic criteria or psychiatric interview using Montgomery–Åsberg Depression Rating Scale (MADRS) Follow-up: 7-20 weeks	Study population 633 per 1000	317 per 1000 (82 to 551)	RR 0.54 (0.31 to 0.93)	179 (3 studies)	⊕⊖⊖⊖ very low[1,3,4]	
	Moderate 625 per 1000	312 per 1000 (81 to 544)				
Depression symptomatology Post-treatment – ITT analysis	Study population 452 per 1000	434 per 1000 (380 to 493)	RR 0.96	1,111 (2 studies)	⊕⊕⊕⊖ moderate[4]	

EPDS≥12 Follow-up: 26-52 weeks	Moderate 494 per 1000	474 per 1000 (415 to 538)	(0.84 to 1.09)		
Depression symptomatology Post-treatment – available case analysis EPDS≥12 Follow-up: 26-52 weeks	Study population 331 per 1000 Moderate 373 per 1000	271 per 1000 (218 to 334) 306 per 1000 (246 to 377)	RR 0.82 (0.66 to 1.01)	885 (2 studies)	⊕⊕⊖⊖ low[1,2,4]
Depression mean scores post-treatment – available case analysis EPDS Follow-up: 20-26 weeks		The mean depression mean scores post-treatment – available case analysis in the intervention groups was 0.34 standard deviations lower (0.55 to 0.14 lower)		375 (2 studies)	⊕⊕⊕⊖ moderate[4]	SMD -0.34 (-0.55 to -0.14)
Depression diagnosis intermediate follow-up (17-24 weeks post-intervention) – ITT analysis SCID Follow-up: mean 20 weeks	Study population 365 per 1000 Moderate 365 per 1000	354 per 1000 (208 to 599) 354 per 1000 (208 to 599)	RR 0.97 (0.57 to 1.64)	100 (1 study)	⊕⊕⊖⊖ low[1,2]	
Depression diagnosis intermediate follow-up (17-24 weeks post-intervention) – available case analysis SCID Follow-up: mean 20 weeks	Study population 312 per 1000 Moderate 313 per 1000	341 per 1000 (191 to 606) 341 per 1000 (191 to 607)	RR 1.09 (0.61 to 1.94)	95 (1 study)	⊕⊕⊖⊖ low[1,2]	
Depression mean scores Intermediate follow-up (17-24 weeks post-intervention) – by intervention EPDS or CES-D Follow-up: 4-12 weeks		The mean depression mean scores intermediate follow-up (17-24 weeks post-intervention) – by intervention in the intervention groups was 0.07 standard deviations lower (0.35 lower to 0.21 higher)		197 (2 studies)	⊕⊕⊕⊖ moderate[5]	SMD -0.07 (-0.35 to 0.21)
Depression mean scores Intermediate follow-up (17-24 weeks post-intervention) – available case analysis EPDS Follow-up: mean 20 weeks		The mean depression mean scores intermediate follow-up (17-24 weeks post-intervention) – available case analysis in the intervention groups was 0.07 standard		94 (1 study)	⊕⊕⊖⊖ low[5]	SMD 0.07 (-0.33 to 0.48)

		deviations higher (0.33 lower to 0.48 higher)			
Depression diagnosis Long Follow-up (25-103 weeks post-intervention) – ITT analysis SCID Follow-up: mean 20 weeks	Study population 250 per 1000	355 per 1000 (192 to 650)	RR 1.42 (0.77 to 2.6)	100 (1 study)	⊕⊕⊖⊖ low[1,2]	
	Moderate 250 per 1000	355 per 1000 (192 to 650)				
Depression diagnosis Long Follow-up (25-103 weeks post-intervention) – available case analysis SCID Follow-up: mean 20 weeks	Study population 188 per 1000	311 per 1000 (150 to 647)	RR 1.66 (0.8 to 3.45)	93 (1 study)	⊕⊕⊖⊖ low[1,2]	
	Moderate 188 per 1000	312 per 1000 (150 to 649)				
Depression symptomatology Long Follow-up (25-103 weeks post-intervention) – ITT analysis GHQ≥12 Follow-up: mean 78 weeks	Study population 651 per 1000	638 per 1000 (567 to 723)	RR 0.98 (0.87 to 1.11)	731 (1 study)	⊕⊕⊕⊖ moderate[4]	
	Moderate 652 per 1000	639 per 1000 (567 to 724)				
Depression symptomatology Long Follow-up (25-103 weeks post-intervention) – available case analysis GHQ≥12 Follow-up: mean 78 weeks	Study population 538 per 1000	516 per 1000 (425 to 618)	RR 0.96 (0.79 to 1.15)	549 (1 study)	⊕⊕⊖⊖ low[1,4]	
	Moderate 538 per 1000	516 per 1000 (425 to 619)				
Depression mean scores Long Follow-up (25-103 weeks post-intervention) – available case analysis EPDS Follow-up: mean 78 weeks		The mean depression mean scores long follow-up (25-103 weeks post-intervention) – available case analysis in the intervention groups was 0.14 standard deviations higher (0.26 lower to 0.55 higher)		92 (1 study)	⊕⊕⊖⊖ low[2,5]	SMD 0.14 (-0.26 to 0.55)
Depression diagnosis Very long Follow-up (>104 weeks post-intervention) – ITT analysis SCID Follow-up: mean 260 weeks	Study population 250 per 1000	458 per 1000 (260 to 805)	RR 1.83 (1.04 to 3.22)	100 (1 study)	⊕⊕⊖⊖ low[1]	
	Moderate 250 per 1000	458 per 1000 (260 to 805)				
Depression diagnosis Very long Follow-up (>104 weeks post-intervention) – available	Study population 243 per 1000	212 per 1000 (90 to 506)	RR 0.87	70 (1 study)	⊕⊕⊖⊖ low[1,2]	

case analysis SCID Follow-up: mean 260 weeks	Moderate 243 per 1000	211 per 1000 (90 to 505)	(0.37 to 2.08)		
Depression mean scores Very long Follow-up (>104 weeks post-intervention) – available case analysis EPDS Follow-up: mean 260 weeks		The mean depression mean scores very long follow-up (>104 weeks post-intervention) – available case analysis in the intervention groups was 0.19 standard deviations lower (0.67 lower to 0.29 higher)	67 (1 study)	⊕⊕⊖⊖ low[2,5]	SMD - 0.19 (-0.67 to 0.29)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total number of events is less than 300 (a threshold rule-of-thumb)
[2] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
[3] There was evidence of moderate to substantial heterogeneity between effect sizes
[4] Papers omit data
[5] Total population size is less than 400 (a threshold rule-of-thumb)

Depression: directive counselling versus treatment as usual

There was low quality, single study (N=146) evidence that directive counselling was more effective than treatment as usual for depression symptomatology (using either ITT or available case methods of analysis) with moderate effects observed on dichotomous measures at endpoint (p=0.002-0.003) and a large effect observed on a continuous measure at long-term follow-up (p=0.0005), although it is important to note that the effects on mean depression symptoms at endpoint (p=0.11) were not statistically or clinically significant (Table 137).

Table 137: Summary of findings table for effects of directive counselling compared with treatment as usual on depression outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Depression: Directive counselling versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Depression symptomatology Post-treatment – ITT analysis BDI≥16 Follow-up: mean 12 weeks	Study population 848 per 1000	611 per 1000 (501 to 747)	RR 0.72 (0.59 to 0.88)	146 (1 study)	⊕⊕⊖⊖ low[1]	
	Moderate 849 per 1000	611 per 1000 (501 to 747)				
Depression symptomatology Post-treatment – available case analysis BDI≥16 Follow-up: mean 12 weeks	Study population 722 per 1000	390 per 1000 (260 to 585)	RR 0.54 (0.36 to 0.81)	90 (1 study)	⊕⊕⊖⊖ low[1]	
	Moderate 722 per 1000	390 per 1000 (260 to 585)				
Depression mean scores post-treatment – available case analysis BDI Follow-up: mean 12 weeks		The mean depression mean scores post-treatment – available case analysis in the intervention groups was 0.42 standard deviations lower (0.95 lower to 0.1 higher)		90 (1 study)	⊕⊕⊖⊖ low[2,3]	SMD -0.42 (-0.95 to 0.1)
Depression mean scores Long Follow-up (25-103 weeks post-intervention) – available case analysis BDI Follow-up: mean 52 weeks		The mean depression mean scores long follow-up (25-103 weeks post-intervention) – available case analysis in the intervention groups was 1.46 standard deviations lower (2.29 to 0.63 lower)		45 (1 study)	⊕⊕⊖⊖ low[2]	SMD -1.46 (-2.29 to -0.63)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total number of events is less than 300 (a threshold rule-of-thumb)
[2] Total population size is less than 400 (a threshold rule-of-thumb)
[3] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Depression: post-miscarriage counselling versus treatment as usual or enhanced treatment as usual

There was no evidence for clinically or statistically significant benefits associated with post-miscarriage counselling on mean depression symptoms at endpoint (ITT [p=0.24] or available case [p=0.52] analysis) or at intermediate (p=0.36) or long (p=0.62) follow-ups (Table 138).

Table 138: Summary of findings table for effects of post-miscarriage counselling compared with treatment as usual on depression outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Depression: post-miscarriage counselling versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Depression mean scores post-treatment – ITT analysis CES-D or HRSD Follow-up: 7-12 weeks		The mean depression mean scores post-treatment – ITT analysis in the intervention groups was 0.17 standard deviations higher (0.12 lower to 0.46 higher)		189 (2 studies)	⊕⊕⊖⊖ low[1]	SMD 0.17 (-0.12 to 0.46)
Depression mean scores post-treatment – available case analysis HRSD or HADS – Depression Follow-up: 2-7 weeks		The mean depression mean scores post-treatment – available case analysis in the intervention groups was 0.14 standard deviations higher (0.29 lower to 0.58 higher)		81 (2 studies)	⊕⊕⊖⊖ low[1,2]	SMD 0.14 (-0.29 to 0.58)

Depression mean scores Intermediate follow-up (17-24 weeks post-intervention) – available case analysis HADS – Depression Follow-up: mean 17 weeks	The mean depression mean scores intermediate follow-up (17-24 weeks post-intervention) – available case analysis in the intervention groups was 0.23 standard deviations lower (0.71 lower to 0.26 higher)	66 (1 study)	⊕⊕⊖⊖ low[1,2]	SMD -0.23 (-0.71 to 0.26)
Depression mean scores Long Follow-up (25-103 weeks post-intervention) – ITT analysis CES-D Follow-up: mean 46 weeks	The mean depression mean scores long follow-up (25-103 weeks post-intervention) – ITT analysis in the intervention groups was 0.08 standard deviations lower (0.38 lower to 0.22 higher)	170 (1 study)	⊕⊕⊖⊖ low[1]	SMD -0.08 (-0.38 to 0.22)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total population size is less than 400 (a threshold rule-of-thumb)
[2] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Depression: post-traumatic birth counselling versus treatment as usual

There was low quality, single study (N=103) evidence for large effects (p=0.008) of post-traumatic birth counselling on depression symptomatology (Table 139).

Table 139: Summary of findings table for effects of post-traumatic birth counselling compared with treatment as usual on depression outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Depression: post-traumatic birth counselling versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Depression symptomatology Post-treatment – ITT analysis EPDS≥12 Follow-up: mean 13 weeks	Study population 321 per 1000	80 per 1000 (29 to 221)	RR 0.25 (0.09 to 0.69)	103 (1 study)	⊕⊕⊖⊖ low[1]	
	Moderate 321 per 1000	80 per 1000 (29 to 221)				
Depression symptomatology Post-treatment – available case analysis EPDS≥12 Follow-up: mean 13 weeks	Study population 321 per 1000	80 per 1000 (29 to 221)	RR 0.25 (0.09 to 0.69)	103 (1 study)	⊕⊕⊖⊖ low[1]	
	Moderate 321 per 1000	80 per 1000 (29 to 221)				

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total number of events is less than 300 (a threshold rule-of-thumb)

Depression: social support versus treatment as usual

There were mixed results for treatment effects on depression outcomes associated with peer-mediated support or support groups (mutual support). There was low to moderate quality evidence from three studies (N=713/807) for moderate benefits of social support on depression symptomatology at endpoint using an ITT (p=0.05) or available case (p <0.0001) data analysis approach (Table 140). However, these effects appeared to be transient as no clinically or statistically significant benefits (p=0.38-0.40) were observed on depression symptomatology at short-term follow-up (9-16 weeks post-intervention). Moreover, there was no evidence for clinically or

statistically significant benefits of social support on depression diagnosis at endpoint using ITT analysis (p=0.52) or for mean depression symptoms at endpoint (p=0.68) or short-term follow-up (p=0.11) and no statistically significant treatment effects on depression diagnosis at endpoint using an available case analysis approach (p=0.18).

Table 140: Summary of findings table for effects of social support compared with treatment as usual on depression outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Depression: Social support versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Depression diagnosis post-treatment – ITT analysis SCID Follow-up: mean 12 weeks	Study population		RR 1.11 (0.81 to 1.52)	701 (1 study)	⊕⊖⊖⊖ very low1,2,3	
	170 per 1000	189 per 1000 (138 to 259)				
	Moderate					
	171 per 1000	190 per 1000 (139 to 260)				
Depression diagnosis post-treatment – available case analysis SCID Follow-up: mean 12 weeks	Study population		RR 0.65 (0.34 to 1.23)	612 (1 study)	⊕⊖⊖⊖ very low1,2,3	
	73 per 1000	47 per 1000 (13 to 83)				
	Moderate					
	73 per 1000	47 per 1000 (13 to 83)				
Depression symptomatology Post-treatment – ITT analysis BDI≥10 or EPDS≥12 Follow-up: 8-14 weeks	Study population		RR 0.69 (0.47 to 1.01)	807 (3 studies)	⊕⊕⊖⊖ low1,2	
	359 per 1000	248 per 1000 (169 to 363)				
	Moderate					
	546 per 1000	377 per 1000 (257 to 551)				
Depression symptomatology Post-treatment – available case analysis BDI≥10 or EPDS≥12 Follow-up: 8-14 weeks	Study population		RR 0.52 (0.39 to 0.7)	713 (3 studies)	⊕⊕⊕⊖ moderate1	
	292 per 1000	152 per 1000 (114 to 205)				
	Moderate					
	524 per 1000	272 per 1000 (204 to 367)				
Depression mean scores post-treatment – available case analysis BDI or EPDS Follow-up: 12-14 weeks		The mean depression mean scores post-treatment – available case analysis in the intervention groups was 0.12 standard deviations lower		723 (3 studies)	⊕⊖⊖⊖ very low2,4	SMD -0.12 (-0.68 to 0.45)

		(0.68 lower to 0.45 higher)			
Depression symptomatology Short Follow-up (9-16 weeks post-intervention) – ITT analysis EPDS≥12 Follow-up: mean 24 weeks	Study population 239 per 1000	267 per 1000 (208 to 344)	RR 1.12 (0.87 to 1.44)	701 (1 study)	⊕⊕⊖⊖ low1,2	
	Moderate 239 per 1000	268 per 1000 (208 to 344)				
Depression symptomatology Short Follow-up (9-16 weeks post-intervention) – available case analysis EPDS≥12 Follow-up: mean 24 weeks	Study population 138 per 1000	115 per 1000 (75 to 174)	RR 0.83 (0.54 to 1.26)	600 (1 study)	⊕⊕⊖⊖ low1,2	
	Moderate 138 per 1000	115 per 1000 (75 to 174)				
Depression mean scores Short Follow-up (9-16 weeks post-intervention) – available case analysis EPDS Follow-up: mean 24 weeks		The mean depression mean scores short follow-up (9-16 weeks post-intervention) – available case analysis in the intervention groups was 0.13 standard deviations lower (0.29 lower to 0.03 higher)		600 (1 study)	⊕⊕⊕⊕ high	SMD -0.13 (-0.29 to 0.03)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total number of events is less than 300 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
3 Papers omit data
4 There was evidence of considerable heterogeneity between effect sizes

Depression: Psychologically (CBT/IPT)-informed psychoeducation versus treatment as usual or enhanced treatment as usual

There was inconsistent evidence for benefits associated with psychologically-informed psychoeducation. There was evidence from up to eight studies (N=985) for

moderate effects of psychoeducation on depression diagnosis at endpoint using an ITT or available case data analysis approach (p=0.10) and at long-term follow-up (25-103 weeks post-intervention) using an available case analysis approach (p=0.06), however, the confidence in these effect estimates is very low due to the 95% CI including both estimates of no effect and estimates of appreciable clinical benefit (Table 141). There was also high quality evidence from five studies (N=1,518) for small to moderate (statistically significant) benefits associated with psychoeducation observed on depression symptomatology (ITT [p=0.0008] and available case [p=0.03] analysis), however, here it is unclear that benefits were clinically meaningful with the treatment effect in the available case analysis falling below the threshold for clinically meaningful benefit. Treatment effects of psychoeducation on mean depression scores at endpoint (although in many cases statistically significant) also failed to reach the threshold for clinically significant benefits at endpoint (using either ITT [p=0.13] or available case [p=0.01] analysis approaches) or at short-term (9-16 week post-intervention) follow-up (with ITT [p=0.005] or available case [p=0.04] analysis) or long-term follow-up (with ITT [p=0.05] or available case [p=0.006] analysis). There was also no evidence for any statistically or clinically significant treatment effects for any outcome measures at intermediate (17-24 weeks post-intervention) follow-up (p=0.38-0.78) or for depression diagnosis at long-term follow-up using an ITT analysis approach (p=0.20).

Table 141: Summary of findings table for effects of psychologically (CBT/IPT)-informed psychoeducation compared with treatment as usual or enhanced treatment as usual on depression outcomes

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Depression: Psychologically (CBT/IPT)-informed psychoeducation versus TAU/Enhanced TAU				
Depression diagnosis post-treatment – ITT analysis MINI, Schedule for Affective Disorders and Schizophrenia (SADS), Maternal Mood Screener (MMS), SCID or Longitudinal Interval Follow-up	Study population		RR 0.67 (0.41 to 1.08)	985 (8 studies)	⊕⊖⊖⊖ very low1,2,3	
	163 per 1000	109 per 1000 (67 to 176)				
	Moderate					
	239 per 1000	160 per 1000 (98 to 258)				

Examination (LIFE) Follow-up: 4-52 weeks						
Depression diagnosis post-treatment – available case analysis SADS, MMS or SCID or LIFE Follow-up: 4-52 weeks	Study population 170 per 1000	71 per 1000 (-31 to 180)	RR 0.50 (0.22 to 1.14)	464 (6 studies)	⊕⊖⊖⊖ very low1,2,3,4	
	Moderate 219 per 1000	92 per 1000 (-39 to 232)				
Depression symptomatology Post-treatment – ITT analysis HSCL: Sum/20 >0.75 depression, EPDS≥13, Leverton Questionnaire (Elliott et al., 2000) ≥12, QIDS ≥11 or BDI: Treatment non-response Follow-up: 4-26 weeks	Study population 351 per 1000	260 per 1000 (218 to 309)	RR 0.74 (0.62 to 0.88)	1,518 (5 studies)	⊕⊕⊕⊕ high	
	Moderate 480 per 1000	355 per 1000 (298 to 422)				
Depression symptomatology Post-treatment – available case analysis HSCL: Sum/20 >0.75 depression, QIDS≥11 or BDI: Treatment non-response Follow-up: 4-26 weeks	Study population 320 per 1000	262 per 1000 (218 to 314)	RR 0.82 (0.68 to 0.98)	997 (3 studies)	⊕⊕⊕⊖ moderate1	
	Moderate 458 per 1000	376 per 1000 (311 to 449)				
Depression mean scores post-treatment – ITT analysis EPDS or CES-D Follow-up: 4-31 weeks		The mean depression mean scores post-treatment – ITT analysis in the intervention groups was 0.25 standard deviations lower (0.58 lower to 0.08 higher)		436 (4 studies)	⊕⊕⊕⊖ moderate4	SMD -0.25 (-0.58 to 0.08)
Depression mean scores post-treatment – available case analysis BDI-II, BDI, EPDS or CES-D Follow-up: 4-31 weeks		The mean depression mean scores post-treatment – available case analysis in the intervention groups was 0.26 standard deviations lower (0.48 to 0.05 lower)		351 (7 studies)	⊕⊕⊕⊖ moderate5	SMD -0.26 (-0.48 to -0.05)
Depression mean scores Short Follow-up (9-16 weeks post-intervention) – ITT analysis		The mean depression mean scores short follow-up (9-16 weeks post-intervention) – ITT analysis in the		235 (2 studies)	⊕⊕⊕⊖ moderate5	SMD -0.37 (-0.63 to -0.11)

EPDS Follow-up: 13-27 weeks		intervention groups was 0.37 standard deviations lower (0.63 to 0.11 lower)				
Depression mean scores Short Follow-up (9-16 weeks post-intervention) – available case analysis EPDS or BDI-II Follow-up: 19-27 weeks		The mean depression mean scores short follow-up (9-16 weeks post-intervention) – available case analysis in the intervention groups was 0.42 standard deviations lower (0.82 to 0.02 lower)		100 (2 studies)	⊕⊖⊖⊖ very low3,5	SMD -0.42 (-0.82 to -0.02)
Depression diagnosis intermediate follow-up (17-24 weeks post-intervention) – ITT analysis MINI, SADS or MMS Follow-up: 6-36 weeks	Study population 113 per 1000	125 per 1000 (85 to 181)	RR 1.1 (0.75 to 1.6)	734 (4 studies)	⊕⊖⊖⊖ very low1,2,3,6	
	Moderate 86 per 1000	95 per 1000 (65 to 138)				
Depression diagnosis intermediate follow-up (17-24 weeks post-intervention) – available case analysis SADS or MMS Follow-up: 26-36 weeks	Study population 128 per 1000	141 per 1000 (74 to 268)	RR 1.1 (0.58 to 2.09)	233 (2 studies)	⊕⊖⊖⊖ very low1,2,3	
	Moderate 77 per 1000	85 per 1000 (45 to 161)				
Depression mean scores Intermediate follow-up (17-24 weeks post-intervention) – ITT analysis EPDS Follow-up: 26-36 weeks		The mean depression mean scores intermediate follow-up (17-24 weeks post-intervention) – ITT analysis in the intervention groups was 0.07 standard deviations lower (0.35 lower to 0.21 higher)		197 (2 studies)	⊕⊕⊖⊖ low5	SMD -0.07 (-0.35 to 0.21)
Depression mean scores Intermediate follow-up (17-24 weeks post-intervention) – available case analysis EPDS Follow-up: mean 36 weeks		The mean depression mean scores intermediate follow-up (17-24 weeks post-intervention) – available case analysis in the intervention groups was 0.28 standard deviations lower (0.89 lower to 0.34 higher)		41 (1 study)	⊕⊖⊖⊖ very low2,3,5,6	SMD -0.28 (-0.89 to 0.34)

Depression diagnosis Long Follow-up (25-103 weeks post-intervention) – ITT analysis MINI, SADS, MMS or SCID Follow-up: 32-75 weeks	Study population 217 per 1000 — 173 per 1000 (121 to 245) Moderate 250 per 1000 — 200 per 1000 (140 to 282)	RR 0.8 (0.56 to 1.13)	812 (5 studies)	⊕⊖⊖⊖ very low1,2,3	
Depression diagnosis Long Follow-up (25-103 weeks post-intervention) – available case analysis SADS, MMS or SCID Follow-up: 32-75 weeks	Study population 227 per 1000 — 136 per 1000 (82 to 233) Moderate 250 per 1000 — 150 per 1000 (90 to 257)	RR 0.6 (0.36 to 1.03)	266 (3 studies)	⊕⊖⊖⊖ very low1,2,3	
Depression mean scores Long Follow-up (25-103 weeks post-intervention) – ITT analysis EPDS Follow-up: 57-75 weeks	The mean depression mean scores long follow-up (25-103 weeks post-intervention) – ITT analysis in the intervention groups was 0.43 standard deviations lower (0.86 lower to 0 higher)		86 (2 studies)	⊕⊕⊖⊖ low5	SMD -0.43 (-0.86 to 0)
Depression mean scores Long Follow-up (25-103 weeks post-intervention) – available case analysis EPDS or BDI-II Follow-up: 32-75 weeks	The mean depression mean scores long follow-up (25-103 weeks post-intervention) – available case analysis in the intervention groups was 0.44 standard deviations lower (0.75 to 0.12 lower)		161 (3 studies)	⊕⊖⊖⊖ very low3,5	SMD -0.44 (-0.75 to -0.12)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total number of events is less than 300 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
3 Papers omit data

4 There was evidence of substantial heterogeneity between effect sizes
5 Total population size is less than 400 (a threshold rule-of-thumb)
6 Risk of bias due to statistically significant group differences at baseline

Depression: Psychologically (CBT/IPT)-informed psychoeducation versus alternative active intervention

There was no evidence that IPT-informed psychoeduation was more effective than non-mental health-focused education and support for treating depression symptomatology (p=0.12; Table 142).

Table 142: Summary of findings table for effects of IPT-informed psychoeducation compared with non-mental health-focused education and support on depression outcomes

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Depression: IPT-informed psychoeducation versus non-mental health-focused education and support				
Depression symptomatology Post-treatment – ITT Analysis EPDS Follow-up: mean 16 weeks	Study population		RR 0.76 (0.53 to 1.07)	38 (1 study)	⊕⊕⊖⊖ low[1,2]	
	882 per 1000	671 per 1000 (468 to 944)				
	Moderate					
	882 per 1000	670 per 1000 (467 to 944)				

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total number of events is less than 300 (a threshold rule-of-thumb)
[2] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Depression: non-mental health-focused education and support versus treatment as usual

There was no evidence for clinically or statistically significant benefits (p=0.07) associated with non-mental health-focused education and support for depression symptomatology (Table 143).

Table 143: Summary of findings table for effects of non-mental health-focused education and support compared with treatment as usual on depression outcomes

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Depression: Non-mental health-focused education and support versus TAU				
Depression symptomatology Post-treatment – ITT analysis HSCL-25 >1.06 Follow-up: mean 12 weeks	Study population		RR 0.91 (0.82 to 1.01)	331 (1 study)	⊕⊕⊕⊖ moderate[1]	
	847 per 1000	770 per 1000 (694 to 855)				
	Moderate					
	847 per 1000	771 per 1000 (695 to 855)				
Depression symptomatology Post-treatment – available case analysis HSCL-25 >1.06 Follow-up: mean 12 weeks	Study population		RR 0.82 (0.67 to 1.01)	188 (1 study)	⊕⊕⊖⊖ low[1,2]	
	725 per 1000	595 per 1000 (486 to 733)				
	Moderate					
	725 per 1000	595 per 1000 (486 to 732)				

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total number of events is less than 300 (a threshold rule-of-thumb)
[2] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Depression: home visits versus treatment as usual or enhanced treatment as usual

There was single study (N=16-18) evidence for large (available case analysis [p=0.19]) to moderate (ITT analysis [p=0.36]) benefits of home visits on depression diagnosis (Table 144). However, confidence in these effect estimates is very low due to the 95% CI including estimates of both no effect and clinically meaningful treatment benefits. Moreover, there was no evidence of clinically or statistically significant treatment effects on depression symptomology (p=0.23-0.24), or clinically significant treatment effects on mean depression symptoms (p=0.008).

Table 144: Summary of findings table for effects of home visits compared with treatment as usual or enhanced treatment as usual on depression outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Depression: Home visits versus TAU/Enhanced TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Depression diagnosis post-treatment – ITT analysis SCID Follow-up: mean 6 weeks	Study population		RR 0.67 (0.28 to 1.58)	18 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	
	667 per 1000	447 per 1000 (187 to 1000)				
	Moderate					
	667 per 1000	447 per 1000 (187 to 1000)				
Depression diagnosis post-treatment – available case analysis SCID Follow-up: mean 6 weeks	Study population		RR 0.43 (0.12 to 1.51)	16 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	
	667 per 1000	287 per 1000 (-173 to 740)				
	Moderate					
	667 per 1000	287 per 1000 (-173 to 740)				
Depression symptomatology Post-treatment – ITT analysis EPDS ≥10/12 or CES-D ≥24 Follow-up: 22-104 weeks	Study population		RR 0.92 (0.8 to 1.06)	985 (3 studies)	⊕⊕⊕⊖ moderate[4]	
	451 per 1000	415 per 1000 (361 to 479)				
	Moderate					
	477 per 1000	439 per 1000 (382 to 506)				
Depression symptomatology Post-treatment –	Study population		RR 0.87 (0.69 to 1.1)	754 (3 studies)	⊕⊖⊖⊖ very low[2,3,4]	
	279 per 1000	243 per 1000 (193 to 307)				

available case analysis EPDS ≥10/12 or CES-D ≥24 Follow-up: 22-104 weeks	Moderate 220 per 1000	191 per 1000 (152 to 242)			
Depression mean scores post-treatment – available case analysis EPDS or CES-D Follow-up: 22-52 weeks		The mean depression mean scores post-treatment – available case analysis in the intervention groups was 0.17 standard deviations lower (0.3 to 0.05 lower)	960 (3 studies)	⊕⊕⊕⊕ high	SMD -0.17 (-0.3 to -0.05)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Risk of bias due to unclear blinding of outcome assessment
[2] Total number of events is less than 300 (a threshold rule-of-thumb)
[3] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
[4] Papers omit data

Depression: mother-infant relationship interventions versus treatment as usual or enhanced treatment as usual

Evidence for treatment effects of mother–infant relationship interventions on depression outcome measures was very inconsistent (Table 145). There was single study (N=92-95) evidence for moderate benefits of a mother–infant relationship intervention on depression diagnosis at endpoint (p=0.10-0.11) and very long-term follow-up (>103 weeks post-intervention) using available case analysis (p=0.42). However, the quality of this evidence was low due to very serious imprecision (with small number of events and 95% CIs including estimates of no effect and clinically meaningful benefit). Conversely, there was single study evidence suggestive of harms associated with mother–infant relationship interventions on depression symptomatology at intermediate (17-24 weeks post-intervention) follow-up (p=0.40-0.42) and depression diagnosis at long-term follow-up (25-103 weeks post-intervention) using available case analysis (p=0.28). However, again the quality of the evidence is low due to very serious imprecision. In addition, low quality evidence from meta-analyses with up to six studies (N=566) provided no evidence for clinically or statistically significant benefits of mother–infant relationship interventions on depression symptomatology at endpoint (p=0.25-0.41), or

depression mean symptoms at endpoint (p=0.93) or long-term follow-up (p=0.61). Single study data for depression diagnosis and depression mean symptoms at intermediate follow-up, depression diagnosis at long-term follow-up (using ITT analysis) or very long-term follow-up (using ITT analysis), and depression mean symptoms at very long-term follow-up also provided no evidence for clinically or statistically significant treatment effects (p=0.49-0.62).

A single study also examined differences between two active intervention arms and found no advantage to video feedback compared with verbal feedback (p=0.38) for effects of mother–infant relationship interventions on mean depression symptoms (Table 146).

Table 145: Summary of findings table for effects of mother–infant relationship interventions compared with treatment as usual or enhanced treatment as usual on depression outcomes

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Depression: Mother-infant relationship interventions versus TAU/Enhanced TAU				
Depression diagnosis post-treatment – ITT analysis SCID Follow-up: mean 20 weeks	Study population		RR 0.72 (0.48 to 1.07)	95 (1 study)	⊕⊕⊖⊖ low[1,2]	
	615 per 1000	443 per 1000 (295 to 658)				
	Moderate					
	615 per 1000	443 per 1000 (295 to 658)				
Depression diagnosis post-treatment – available case analysis SCID Follow-up: mean 20 weeks	Study population		RR 0.71 (0.47 to 1.08)	92 (1 study)	⊕⊕⊖⊖ low[1,2]	
	600 per 1000	426 per 1000 (228 to 630)				
	Moderate					
	600 per 1000	426 per 1000 (228 to 630)				
Depression symptomatology Post-treatment – ITT analysis EPDS: Treatment non-response (reliable change index-no improvement)/EPDS ≥12 or CES-D ≥16 Follow-up: 5-26 weeks	Study population		RR 0.87 (0.69 to 1.1)	396 (3 studies)	⊕⊕⊖⊖ low[1,2]	
	565 per 1000	492 per 1000 (390 to 622)				
	Moderate					
	717 per 1000	624 per 1000 (495 to 789)				

Depression symptomatology Post-treatment – available case analysis EPDS: Treatment non-response (reliable change index-no improvement)/EPDS ≥12 or CES-D ≥16 Follow-up: 5-26 weeks	Study population 379 per 1000 Moderate 472 per 1000	322 per 1000 (220 to 473) 401 per 1000 (274 to 590)	RR 0.85 (0.58 to 1.25)	288 (3 studies)	⊕⊕⊖⊖ low[1,2]	
Depression mean scores post-treatment – available case EPDS, BDI, BDI-II or CES-D Follow-up: 5-28 weeks		The mean depression mean scores post-treatment – available case in the intervention groups was 0.02 standard deviations higher (0.38 lower to 0.41 higher)		566 (6 studies)	⊕⊕⊖⊖ low[3]	SMD 0.02 (-0.38 to 0.41)
Depression diagnosis intermediate follow-up (17-24 weeks post-intervention) – ITT analysis SCID Follow-up: mean 39 weeks	Study population 365 per 1000 Moderate 365 per 1000	303 per 1000 (168 to 541) 303 per 1000 (168 to 540)	RR 0.83 (0.46 to 1.48)	95 (1 study)	⊕⊕⊖⊖ low[1,2]	
Depression diagnosis intermediate follow-up (17-24 weeks post-intervention) – available case analysis SCID Follow-up: mean 39 weeks	Study population 312 per 1000 Moderate 313 per 1000	250 per 1000 (125 to 494) 250 per 1000 (125 to 495)	RR 0.8 (0.4 to 1.58)	88 (1 study)	⊕⊕⊖⊖ low[1,2]	
Depression symptomatology Intermediate follow-up (17-24 weeks post-intervention) – ITT analysis EPDS≥12 Follow-up: mean 25 weeks	Study population 262 per 1000 Moderate 262 per 1000	333 per 1000 (191 to 580) 333 per 1000 (191 to 579)	RR 1.27 (0.73 to 2.21)	121 (1 study)	⊕⊕⊖⊖ low[1,2]	
Depression symptomatology Intermediate follow-up (17-24 weeks post-intervention) – available case analysis EPDS ≥12 Follow-up: mean 25 weeks	Study population 80 per 1000 Moderate 80 per 1000	130 per 1000 (39 to 433) 130 per 1000 (39 to 433)	RR 1.63 (0.49 to 5.41)	96 (1 study)	⊕⊕⊖⊖ low[1,2]	

Depression mean scores Intermediate follow-up (17-24 weeks post-intervention) – available case analysis EPDS Follow-up: mean 39 weeks		The mean depression mean scores intermediate follow-up (17-24 weeks post-intervention) – available case analysis in the intervention groups was 0.11 standard deviations lower (0.53 lower to 0.31 higher)		88 (1 study)	⊕⊕⊖⊖ low[2,4]	SMD -0.11 (-0.53 to 0.31)
Depression diagnosis Long Follow-up (25-103 weeks post-intervention) – ITT analysis SCID Follow-up: mean 78 weeks	Study population		RR 1.21 (0.63 to 2.33)	95 (1 study)	⊕⊕⊖⊖ low[1,2]	
	250 per 1000	302 per 1000 (157 to 582)				
	Moderate					
	250 per 1000	302 per 1000 (157 to 582)				
Depression diagnosis Long Follow-up (25-103 weeks post-intervention) – available case analysis SCID Follow-up: mean 78 weeks	Study population		RR 1.52 (0.71 to 3.25)	90 (1 study)	⊕⊕⊖⊖ low[1,2]	
	188 per 1000	285 per 1000 (133 to 609)				
	Moderate					
	188 per 1000	286 per 1000 (133 to 611)				
Depression mean scores Long Follow-up (25-103 weeks post-intervention) – available case analysis EPDS or BDI Follow-up: 57-78 weeks		The mean depression mean scores long follow-up (25-103 weeks post-intervention) – available case analysis in the intervention groups was 0.08 standard deviations higher (0.23 lower to 0.39 higher)		161 (2 studies)	⊕⊕⊖⊖ low[4]	SMD 0.08 (-0.23 to 0.39)
Depression diagnosis Very long Follow-up (≥104 weeks post-intervention) – ITT analysis SCID Follow-up: mean 260 weeks	Study population		RR 1.21 (0.63 to 2.33)	95 (1 study)	⊕⊕⊖⊖ low[1,2]	
	250 per 1000	302 per 1000 (157 to 582)				
	Moderate					
	250 per 1000	302 per 1000 (157 to 582)				
Depression diagnosis Very long Follow-up (≥104 weeks post-intervention) – available	Study population		RR 0.69 (0.27 to 1.73)	73 (1 study)	⊕⊕⊖⊖ low[1,2]	
	243 per 1000	168 per 1000 (66 to 421)				
	Moderate					

case analysis SCID Follow-up: mean 260 weeks	243 per 1000	168 per 1000 (66 to 420)			
Depression mean scores Very long Follow-up (≥104 weeks post-intervention) – available case analysis EPDS Follow-up: mean 260 weeks		The mean depression mean scores very long follow-up (≥104 weeks post-intervention) – available case analysis in the intervention groups was 0.17 standard deviations lower (0.66 lower to 0.32 higher)	65 (1 study)	⊕⊕⊖⊖ low[2,4]	SMD -0.17 (-0.66 to 0.32)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total number of events is less than 300 (a threshold rule-of-thumb)
[2] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
[3] There was evidence of considerable heterogeneity between effect sizes
[4] Total population size is less than 400 (a threshold rule-of-thumb)

Table 146: Summary of findings table for effects of mother–infant relationship intervention with video feedback compared with mother–infant relationship intervention with verbal feedback on depression outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Depression: Mother-infant relationship intervention with video feedback versus mother–infant relationship intervention with verbal feedback	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Depression mean scores		The mean depression mean scores post-treatment –		37 (1 study)	⊕⊕⊖⊖ low[1,2]	SMD 0.29 (-0.36 to 0.94)

post-treatment – available case analysis EPDS Follow-up: mean 3 weeks	available case analysis in the intervention groups was 0.29 standard deviations higher (0.36 lower to 0.94 higher)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total population size is less than 400 (a threshold rule-of-thumb)
[2] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Depression: co-parenting intervention versus enhanced treatment as usual

There was single study (N=29) evidence for a moderate effect of a co-parenting intervention on depression diagnosis (p=0.12). However, confidence in this effect estimate was very low due to very serious imprecision (small number of events and a large 95% CI encompassing no effects and appreciable benefits). In addition, the same study showed no evidence for statistically or clinically significant benefits of a co-parenting intervention on mean depression symptoms (p=0.23; Table 147).

Table 147: Summary of findings table for effects of co-parenting intervention compared with enhanced treatment as usual on depression outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Depression: Co-parenting intervention versus Enhanced TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Depression diagnosis post-treatment – ITT analysis MINI Follow-up: mean 6 weeks	Study population		RR 0.51 (0.22 to 1.18)	29 (1 study)	⊕⊖⊖⊖ very low1,2,3	
	615 per 1000	314 per 1000 (135 to 726)				
	Moderate					
	615 per 1000	314 per 1000 (135 to 726)				
	Study population					

Depression diagnosis post-treatment – available case analysis MINI Follow-up: mean 6 weeks	615 per 1000	314 per 1000 (-37 to 665)	RR 0.51 (0.22 to 1.18)	29 (1 study)	⊕⊖⊖⊖ very low1,2,3	
	Moderate					
	615 per 1000	314 per 1000 (-37 to 664)				
Depression mean scores post-treatment – available case analysis EPDS Follow-up: mean 6 weeks		The mean depression mean scores post-treatment – available case analysis in the intervention groups was 0.47 standard deviations lower (1.22 lower to 0.29 higher)		28 (1 study)	⊕⊖⊖⊖ very low1,3,4	SMD -0.47 (-1.22 to 0.29)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Risk of bias as blinding of outcome assessment was unclear
2 Total number of events is less than 300 (a threshold rule-of-thumb)
3 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
4 Total population size is less than 400 (a threshold rule-of-thumb)

Depression: infant sleep training (controlled crying) versus treatment as usual or enhanced treatment as usual

There was low quality single study (N=272) evidence for moderate effects of infant sleep training (controlled crying) on maternal depression symptomatology (p=0.03). There was also low to moderate quality evidence from up to two studies (N=184-272) for statistically significant benefits of controlled crying on mean depression symptoms at endpoint or first measurement, short-term follow-up, and long-term follow-up (p=0.03-0.001), however, these effects were small and below the threshold for appreciable clinical benefit (Table 148).

Table 148: Summary of findings table for effects of infant sleep training (controlled crying) compared with treatment as usual or enhanced treatment as usual on depression outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Depression: Infant sleep training (controlled crying) versus TAU/Enhanced TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Depression symptomatology Post-treatment – available case analysis EPDS >9 Follow-up: mean 74 weeks	Study population 264 per 1000 Moderate 264 per 1000	153 per 1000 (95 to 248) 153 per 1000 (95 to 248)	RR 0.58 (0.36 to 0.94)	272 (1 study)	⊕⊕⊖⊖ low[1]	
Depression mean scores post-treatment – available case analysis EPDS change score or score at endpoint Follow-up: 9-13 weeks		The mean depression mean scores post-treatment – available case analysis in the intervention groups was 0.47 standard deviations lower (0.76 to 0.18 lower)		189 (2 studies)	⊕⊕⊖⊖ low[2]	SMD -0.47 (-0.76 to -0.18)
Depression mean scores Short Follow-up (9-16 weeks post-intervention)- Available case analysis EPDS change score or score at endpoint Follow-up: 17-22 weeks		The mean depression mean scores short follow-up (9-16 weeks post-intervention)- available case analysis in the intervention groups was 0.4 standard deviations lower (0.7 to 0.11 lower)		184 (2 studies)	⊕⊕⊖⊖ low[2]	SMD -0.4 (-0.7 to -0.11)
Depression mean scores Long Follow-up (25-103 weeks post-intervention) – available case analysis EPDS Follow-up: mean 74 weeks		The mean depression mean scores long follow-up (25-103 weeks post-intervention) – available case analysis in the intervention groups was 0.26 standard		272 (1 study)	⊕⊕⊕⊖ moderate[2]	SMD -0.26 (-0.5 to -0.02)

	deviations lower (0.5 to 0.02 lower)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total number of events is less than 300 (a threshold rule-of-thumb)
[2] Total population size is less than 400 (a threshold rule-of-thumb)

Depression: music therapy during birth versus treatment as usual

There was low quality, single study (N=141) evidence for large effects of music therapy during birth on depression symptomatology using available case analysis (p=0.04), moderate effects on depression symptomatology using ITT analysis (p=0.07) and small effects on mean depression symptoms immediately post-birth (p=0.03). However, there was serious imprecision across all outcome measures due to the low number of events or small sample size and/or large 95% CIs encompassing estimates of no effect and appreciable benefit (Table 149).

Table 149: Summary of findings table for effects of music therapy during birth compared with treatment as usual on depression outcomes

Outcomes	**Illustrative comparative risks* (95% CI)**		**Relative effect (95% CI)**	**No. of Participants (studies)**	**Quality of the evidence (GRADE)**	**Comments**
	Assumed risk	**Corresponding risk**				
	Control	**Depression: Music therapy during birth versus TAU**				
Depression symptomatology Post-treatment – ITT analysis EPDS ≥13 Follow-up: mean 3 weeks	Study population		RR 0.57 (0.31 to 1.05)	161 (1 study)	⊕⊕⊖⊖ low[1,2]	
	284 per 1000	162 per 1000 (88 to 298)				
	Moderate					
	284 per 1000	162 per 1000 (88 to 298)				
Depression symptomatology Post-treatment – available case	Study population		RR 0.33 (0.11 to 0.97)	141 (1 study)	⊕⊕⊖⊖ low[1]	
	171 per 1000	57 per 1000 (19 to 166)				
	Moderate					

analysis EPDS ≥13 Follow-up: mean 3 weeks	171 per 1000	56 per 1000 (19 to 166)			
Depression mean scores post-treatment – available case analysis EPDS Follow-up: mean 3 weeks		The mean depression mean scores post-treatment – available case analysis in the intervention groups was 0.37 standard deviations lower (0.71 to 0.04 lower)	141 (1 study)	⊕⊕⊖⊖ low[3]	SMD -0.37 (-0.71 to -0.04)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total number of events is less than 300 (a threshold rule-of-thumb)
[2] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
[3] Total population size is less than 400 (a threshold rule-of-thumb)

Depression: Psychosomatic interventions versus treatment as usual

There was no evidence that psychosomatic interventions conferred appreciable and clinically meaningful benefits on depression symptomatology (p=0.04-0.18) or mean depression symptoms (p=0.22; Table 150).

Table 150: Summary of findings table for effects of psychosomatic intervention compared with treatment as usual on depression outcomes

Outcomes	**Illustrative comparative risks* (95% CI)**		**Relative effect (95% CI)**	**No. of Participants (studies)**	**Quality of the evidence (GRADE)**	**Comments**
	Assumed risk	**Corresponding risk**				
	Control	**Depression: Psychosomatic intervention versus TAU**				
Depression symptomatology Post-treatment – ITT	Study population		**RR 0.77** (0.6 to 0.99)	184 (1 study)	⊕⊖⊖⊖ **very low**[1,2]	
	663 per 1000	511 per 1000 (398 to 656)				

analysis EPDS ≥12 Follow-up: mean 34 weeks	Moderate 663 per 1000	511 per 1000 (398 to 656)				
Depression symptomatology Post-treatment – available case analysis EPDS ≥12 Follow-up: mean 34 weeks	Study population 466 per 1000 Moderate 466 per 1000	349 per 1000 (228 to 531) 349 per 1000 (228 to 531)	**RR 0.75** (0.49 to 1.14)	127 (1 study)	⊕⊖⊖⊖ **very low**[1,2,3]	
Depression mean scores post-treatment – available case analysis HADS – Depression or EPDS Follow-up: 34-52 weeks		The mean depression mean scores post-treatment – available case analysis in the intervention groups was 0.21 standard deviations lower (0.54 lower to 0.13 higher)		171 (2 studies)	⊕⊖⊖⊖ **very low**[1,3,4]	SMD -0.21 (-0.54 to 0.13)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Risk of attrition bias due to statistically significant higher drop-out in the control group
[2] Total number of events is less than 300 (a threshold rule-of-thumb)
[3] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
[4] Total population size is less than 400 (a threshold rule-of-thumb)

Depression: mindfulness training versus treatment as usual or enhanced treatment as usual

There was no evidence for statistically or clinically significant benefits associated with mindfulness training on depression mean symptoms (p=0.72) or negative affect mean scores (p=0.38; Table 151).

Table 151: Summary of findings table for effects of mindfulness training compared with treatment as usual or enhanced treatment as usual on depression outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Depression: Mindfulness training versus Enhanced TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Depression mean scores post-treatment – available case analysis CES-D Follow-up: mean 10 weeks		The mean depression mean scores post-treatment – available case analysis in the intervention groups was 0.13 standard deviations lower (0.85 lower to 0.58 higher)		31 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	SMD -0.13 (-0.85 to 0.58)
Negative affect mean scores post-treatment – available case analysis Positive and Negative Affect Schedule-Extended: Negative affect Follow-up: mean 10 weeks		The mean negative affect mean scores post-treatment – available case analysis in the intervention groups was 0.32 standard deviations lower (1.04 lower to 0.4 higher)		31 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	SMD -0.32 (-1.04 to 0.4)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total population size is less than 400 (a threshold rule-of-thumb)
[2] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
[3] Paper omits data

Depression: combined social support and physical exercise versus enhanced treatment as usual

There was single study (N=20) evidence for large benefits of a combined informal social support group and pram walking exercise programme on depression symptomatology (p=0.05) and mean depression symptoms (p=0.002). However, confidence in these effect estimates is low due to the extremely low event rate and very small sample size, and in the case of the depression symptomatology outcome measure the 95% CI includes both no effect and appreciable benefit (Table 152).

Table 152: Summary of findings table for effects of combined social support and physical exercise compared with enhanced treatment as usual on depression outcomes

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Depression: Combined social support and physical exercise versus enhanced TAU				
Depression symptomatology Post-treatment – ITT analysis EPDS ≥12 Follow-up: mean 12 weeks	Study population		RR 0.07 (0 to 1.03)	20 (1 study)	⊕⊕⊖⊖ low1,2	
	700 per 1000	49 per 1000 (0 to 721)				
	Moderate					
	700 per 1000	49 per 1000 (0 to 721)				
Depression symptomatology Post-treatment – available case analysis EPDS ≥12 Follow-up: mean 12 weeks	Study population		RR 0.07 (0 to 1.03)	20 (1 study)	⊕⊕⊖⊖ low1,2	
	700 per 1000	49 per 1000 (0 to 721)				
	Moderate					
	700 per 1000	49 per 1000 (0 to 721)				
Depression mean symptoms Post-treatment – ITT analysis EPDS Follow-up: mean 12 weeks		The mean depression mean symptoms post-treatment – ITT analysis in the intervention groups was 1.64 standard deviations lower (2.68 to 0.59 lower)		20 (1 study)	⊕⊕⊖⊖ low3	SMD -1.64 (-2.68 to -0.59)
Depression mean symptoms Post-treatment – available case analysis EPDS Follow-up: mean 12 weeks		The mean depression mean symptoms post-treatment – available case analysis in the intervention groups was 1.64 standard deviations lower (2.68 to 0.59 lower)		20 (1 study)	⊕⊕⊖⊖ low3	SMD -1.64 (-2.68 to -0.59)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total number of events is less than 300 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
3 Total population size is less than 400 (a threshold rule-of-thumb)

Depression: social support versus physical exercise

In order to tease apart the combined psychosocial and physical intervention effect (discussed above), the same researchers compared social support and physical exercise in a head-to-head trial and provided single study (N=20) evidence for a large effect of social support (social support group) relative to physical exercise (pram walking exercise programme) on depression mean symptoms (p=0.03). However, confidence in this effect estimate was low due to imprecision as a result of the very small sample size (Table 153).

Table 153: Summary of findings table for effects of social support compared with physical exercise on depression outcomes

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Depression: Social support versus physical exercise				
Depression mean symptoms Post-Treatment – available case analysis EPDS Follow-up: mean 12 weeks		The depression mean symptoms post-treatment – available case analysis in the intervention groups was 1.09 standard deviations lower (2.07 to 0.11 lower)		19 (1 study)	⊕⊕⊖⊖ low[1]	SMD -1.09 (-2.07 to -0.11)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.

Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total population size is less than 400 (a threshold rule-of-thumb)

7.5.5 Clinical evidence for effects on anxiety outcomes (by intervention)

Summary of findings can be found in the tables presented in this section. The full GRADE evidence profiles and associated forest plots can be found in Appendix 22 and Appendix 19, respectively.

Anxiety: structured psychological interventions (CBT or IPT) versus treatment as usual or enhanced treatment as usual

There was low quality, single study (N=53) evidence for a large effect of a structured psychological intervention on mean state anxiety symptoms (using an ITT analysis approach [p <0.0001]). However, the only meta-analysis possible (two studies, N=315) revealed no evidence for clinically significant benefits (although differences were statistically significant) associated with mean state anxiety symptoms (p=0.002), and the small benefit for trait anxiety symptoms found in a single study analysis also failed to reach the threshold for appreciable benefit despite meeting statistical significance criteria (p=0.002; Table 154).

Table 154: Summary of findings table for effects of structured psychological interventions (CBT or IPT) compared with treatment as usual or enhanced treatment as usual on anxiety outcomes

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Anxiety: Structured psychological interventions versus TAU/Enhanced TAU				
Anxiety mean scores post-treatment – ITT analysis Beck Anxiety Inventory (BAI)		The mean anxiety mean scores post-treatment – ITT analysis in the intervention groups was 1.34 standard deviations lower (1.94 to 0.74 lower)		53 (1 study)	⊕⊕⊖⊖ low1	SMD -1.34 (-1.94 to -0.74)

Follow-up: mean 44 weeks				
Anxiety mean scores post-treatment – available case analysis BAI or STAI-S Follow-up: 12-26 weeks	The mean anxiety mean scores post-treatment – available case analysis in the intervention groups was 0.35 standard deviations lower (0.58 to 0.13 lower)	315 (2 studies)	⊕⊕⊖⊖ low1,2	SMD -0.35 (-0.58 to -0.13)
Trait anxiety mean scores post-treatment – available case analysis STAI-T Follow-up: mean 26 weeks	The mean trait anxiety mean scores post-treatment – available case analysis in the intervention groups was 0.38 standard deviations lower (0.62 to 0.13 lower)	263 (1 study)	⊕⊕⊖⊖ low1,2	SMD -0.38 (-0.62 to -0.13)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total population size is less than 400 (a threshold rule-of-thumb)
2 Papers omit data

Anxiety: structured psychological interventions (CBT or IPT) versus alternative active intervention

There was no evidence for a clinically or statistically significant benefit of CBT relative to RCT on mean anxiety symptoms (p=0.31; Table 155).

Table 155: Summary of findings table for effects of CBT compared with relational constructivist therapy on anxiety outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	 Corresponding risk Anxiety: CBT versus relational constructivist therapy	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Anxiety mean scores post-treatment – available case analysis BAI		The mean anxiety mean scores post-treatment – available case analysis in the intervention groups was 0.26 standard deviations higher (0.25 lower to 0.77 higher)		60 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	SMD 0.26 (-0.25 to 0.77)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total population size is less than 400 (a threshold rule-of-thumb)
[2] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
[3] Papers omit data

There was no evidence for a clinically or statistically significant benefit associated with IPT relative to a support group for treating mean anxiety symptoms (p=0.11; Table 156).

Table 156: Summary of findings table for effects of IPT compared with support group on anxiety outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	 Corresponding risk Anxiety: IPT versus support group	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Anxiety mean scores post-treatment – available case STAI-S Follow-up: mean 12 weeks		The mean anxiety mean scores post-treatment – available case in the intervention groups was 0.48 standard deviations lower (1.09 lower to 0.12 higher)		44 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	SMD -0.48 (-1.09 to 0.12)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Risk of bias as statistically significant group differences at baseline
[2] Total population size is less than 400 (a threshold rule-of-thumb)
[3] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Anxiety: facilitated self-help versus treatment as usual

There was very low quality, single study (N=59-143) evidence (using both available case and ITT data analysis methods) for moderate to large benefits of facilitated self-help relative to treatment as usual for treating anxiety symptomatology (p=0.02-0.03) and for mean anxiety symptoms (p=0.06; Table 157).

Table 157: Summary of findings table for effects of facilitated self-help compared with treatment as usual on anxiety outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	 Corresponding risk Anxiety: Facilitated self-help versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Anxiety symptomatology Post-treatment – ITT analysis Depression Anxiety Stress Scale (DASS): Anxiety ≥8 Follow-up: mean 20 weeks	Study population		RR 0.67 (0.47 to 0.96)	143 (1 study)	⊕⊖⊖⊖ very low1,2	
	569 per 1000	382 per 1000 (268 to 547)				
	Moderate					
	569 per 1000	381 per 1000 (267 to 546)				
Anxiety symptomatology Post-treatment – available case analysis Depression Anxiety Stress Scale (DASS): Anxiety ≥8 Follow-up: mean 20 weeks	Study population		RR 0.24 (0.07 to 0.81)	89 (1 study)	⊕⊖⊖⊖ very low1,2	
	262 per 1000	63 per 1000 (18 to 212)				
	Moderate					
	262 per 1000	63 per 1000 (18 to 212)				
Anxiety mean scores post-treatment – available case analysis Generalised Anxiety Disorder Assessment (GAD-7) Follow-up: mean 17 weeks		The mean anxiety mean scores post-treatment – available case analysis in the intervention groups was 0.5 standard deviations lower (1.02 lower to 0.02 higher)		59 (1 study)	⊕⊖⊖⊖ very low2,3,4	SMD -0.5 (-1.02 to 0.02)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total number of events is less than 300 (a threshold rule-of-thumb)
2 Paper omits data

3 Total population size is less than 400 (a threshold rule-of-thumb)
4 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Anxiety: post-miscarriage self-help versus treatment as usual

There was no evidence for statistically or clinically significant benefits of post-miscarriage self-help on anxiety symptomatology (p=0.35-0.71) or mean symptoms (p=0.33; Table 158).

Table 158: Summary of findings table for effects of post-miscarriage self-help compared with treatment as usual on anxiety outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	 Corresponding risk Anxiety: post-miscarriage self-help versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Anxiety symptomatology Post-treatment – ITT analysis BSI: Anxiety (Treatment non-response: reliable change index) Follow-up: mean 5 weeks	Study population		RR 0.95 (0.71 to 1.26)	78 (1 study)	⊕⊕⊖⊖ low1,2	
	727 per 1000	691 per 1000 (516 to 916)				
	Moderate					
	727 per 1000	691 per 1000 (516 to 916)				
Anxiety symptomatology Post-treatment – available case analysis BSI: Anxiety (Treatment non-response: reliable change index) Follow-up: mean 5 weeks	Study population		RR 0.83 (0.56 to 1.23)	59 (1 study)	⊕⊕⊖⊖ low1,2	
	692 per 1000	575 per 1000 (388 to 852)				
	Moderate					
	692 per 1000	574 per 1000 (388 to 851)				
Anxiety mean scores post-treatment – ITT analysis BSI: Anxiety Follow-up: mean 5 weeks		The mean anxiety mean scores post-treatment – ITT analysis in the intervention groups was 0.23 standard deviations lower		78 (1 study)	⊕⊕⊖⊖ low2,3	SMD -0.23 (-0.68 to 0.23)

	(0.68 lower to 0.23 higher)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total number of events is less than 300 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
3 Total population size is less than 400 (a threshold rule-of-thumb)

Anxiety: listening visits versus treatment as usual

There was low quality single study (N=254-260) evidence for statistically significant effects of listening visits on mean state (p=0.02) and trait (p=0.04) anxiety symptoms (Table 159). However, these effects were small and failed to reach a threshold indicative of clinically significant treatment benefits. In addition, the confidence in the effect estimates was low due to small sample size and selective outcome reporting.

Table 159: Summary of findings table for effects of listening visits compared with treatment as usual on anxiety outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	 Corresponding risk Anxiety: Listening visits versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Anxiety mean scores post-treatment – available case analysis STAI-S Follow-up: mean 26 weeks		The mean anxiety mean scores post-treatment – available case analysis in the intervention groups was 0.29 standard deviations lower (0.53 to 0.04 lower)		260 (1 study)	⊕⊕⊖⊖ low1,2	SMD -0.29 (-0.53 to -0.04)
Trait anxiety mean scores post-treatment – available case		The mean trait anxiety mean scores post-treatment – available case analysis in the		254 (1 study)	⊕⊕⊖⊖ low1,2	SMD -0.26 (-0.51 to -0.02)

analysis STAI-T Follow-up: mean 26 weeks	intervention groups was 0.26 standard deviations lower (0.51 to 0.02 lower)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total population size is less than 400 (a threshold rule-of-thumb)
2 Paper omits data

Anxiety: directive counselling versus treatment as usual

There was low quality single study (N=90) evidence for moderate effects of directive counselling on mean anxiety symptoms (p=0.04) using an available case analysis approach (Table 160).

Table 160: Summary of findings table for effects of directive counselling compared with treatment as usual on anxiety outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	 Corresponding risk Anxiety: Directive counselling versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Anxiety mean scores post-treatment – available case analysis BAI Follow-up: mean 12 weeks		The mean anxiety mean scores post-treatment – available case analysis in the intervention groups was 0.56 standard deviations lower (1.09 to 0.04 lower)		90 (1 study)	⊕⊕⊖⊖ low[1]	SMD -0.56 (-1.09 to -0.04)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the

estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total population size is less than 400 (a threshold rule-of-thumb)

Anxiety: post-miscarriage counselling versus treatment as usual or enhanced treatment as usual

There was no evidence for statistically or clinically significant benefits of post-miscarriage counselling on anxiety mean scores at endpoint (p=0.67) or at intermediate follow-up (p=0.21; Table 161).

Table 161: Summary of findings table for effects of post-miscarriage counselling compared with treatment as usual on anxiety outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Anxiety: post-miscarriage counselling versus Enhanced TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Anxiety mean scores post-treatment – available case analysis HADS – Anxiety Follow-up: mean 2 weeks		The mean anxiety mean scores post-treatment – available case analysis in the intervention groups was 0.11 standard deviations higher (0.38 lower to 0.59 higher)		66 (1 study)	⊕⊕⊖⊖ low1,2	SMD 0.11 (-0.38 to 0.59)
Anxiety mean scores Intermediate follow-up (17-24 weeks post-intervention) – available case analysis HADS – Anxiety Follow-up: mean 17 weeks		The mean anxiety mean scores intermediate follow-up (17-24 weeks post-intervention) – available case analysis in the intervention groups was 0.31 standard deviations lower (0.8 lower to 0.17 higher)		66 (1 study)	⊕⊕⊖⊖ low1,2	SMD -0.31 (-0.8 to 0.17)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the

estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total population size is less than 400 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Anxiety: post-traumatic birth counselling versus treatment as usual

There was single study (N=103) evidence for a large effect of post-traumatic birth counselling on anxiety symptomatology (p=0.10). However, confidence that this is a true measure of the effect is low due to the low number of events and the fact that the 95% CI crosses both the line of no effect and the measure of appreciable benefit (Table 162).

Table 162: Summary of findings table for effects of post-traumatic birth counselling compared with treatment as usual on anxiety outcomes

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Anxiety: post-traumatic birth counselling versus TAU				
Anxiety symptomatology Post-treatment – ITT analysis Depression Anxiety Stress Scale (DASS): Anxiety >9 Follow-up: mean 13 weeks	Study population		**RR 0.18** (0.02 to 1.42)	103 (1 study)	⊕⊕⊖⊖ **low**[1,2]	
	113 per 1000	20 per 1000 (2 to 161)				
	Moderate					
	113 per 1000	20 per 1000 (2 to 160)				
Anxiety symptomatology Post-treatment – available case analysis Depression Anxiety Stress Scale (DASS): Anxiety >9 Follow-up: mean 13 weeks	Study population		**RR 0.18** (0.02 to 1.42)	103 (1 study)	⊕⊕⊖⊖ **low**[1,2]	
	113 per 1000	20 per 1000 (2 to 161)				
	Moderate					
	113 per 1000	20 per 1000 (2 to 160)				

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the

estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total number of events is less than 300 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Anxiety: social support versus treatment as usual

There was no evidence for clinically or statistically significant benefits of social support on anxiety symptomatology (p=0.05-0.47) or anxiety mean symptoms (p=0.08-0.42; Table 163).

Table 163: Summary of findings table for effects of social support compared with treatment as usual on anxiety outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk / Corresponding risk Control / Anxiety: Social support versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Anxiety symptomatology Post-treatment – ITT analysis STAI-S >44 Follow-up: mean 12 weeks	Study population 349 per 1000 / 325 per 1000 (262 to 398) Moderate 349 per 1000 / 325 per 1000 (262 to 398)	**RR 0.93** (0.75 to 1.14)	701 (1 study)	⊕⊕⊖⊖ **low**[1,2]	
Anxiety symptomatology Post-treatment – available case analysis State-Trait Anxiety Inventory (STAI)-State>44 Follow-up: mean 12 weeks	Study population 273 per 1000 / 205 per 1000 (153 to 273) Moderate 273 per 1000 / 205 per 1000 (153 to 273)	**RR 0.75** (0.56 to 1)	612 (1 study)	⊕⊖⊖⊖ **very low**[1,2]	
Anxiety mean scores post-treatment – available case analysis STAI-S Follow-up: mean 12 weeks	The mean anxiety mean scores post-treatment – available case analysis in the intervention groups was 0.14 standard deviations lower		612 (1 study)	⊕⊕⊕⊖ **moderate**[2]	SMD -0.14 (-0.3 to 0.02)

	(0.3 lower to 0.02 higher)			
Anxiety mean scores Short follow-up (9-16 weeks post-intervention) – available case analysis STAI-S Follow-up: mean 24 weeks	The mean anxiety mean scores short follow-up (9-16 weeks post-intervention) – available case analysis in the intervention groups was 0.07 standard deviations lower (0.23 lower to 0.09 higher)	600 (1 study)	⊕⊕⊕⊖ **moderate**[2]	SMD -0.07 (-0.23 to 0.09)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total number of events is less than 300 (a threshold rule-of-thumb)
[2] Paper omits data

Anxiety: psychologically (CBT/IPT)-informed psychoeducation versus treatment as usual or enhanced treatment as usual

There was no evidence for statistically or clinically significant benefits of psychologically-informed psychoeducation for anxiety diagnosis at endpoint (p=0.58-0.89) or at long-term follow-up (p=0.99; Table 164).

Table 164: Summary of findings table for effects of psychologically (CBT/IPT)-informed psychoeducation compared with treatment as usual or enhanced treatment as usual on anxiety outcomes

Outcomes	**Illustrative comparative risks* (95% CI)**		**Relative effect (95% CI)**	**No. of Participants (studies)**	**Quality of the evidence (GRADE)**	**Comments**
	Assumed risk	**Corresponding risk**				
	Control	**Anxiety: Psychologically (CBT/IPT)-informed psychoeducation versus TAU/Enhanced TAU**				
	Study population					

Anxiety diagnosis post-treatment – ITT analysis MINI or SADS Follow-up: 9-52 weeks	136 per 1000	132 per 1000 (83 to 209)	RR 0.97 (0.61 to 1.54)	476 (2 studies)	⊕⊖⊖⊖ very low1,2,3,4
	Moderate				
	138 per 1000	134 per 1000 (84 to 213)			
Anxiety diagnosis post-treatment – available case analysis SADS Follow-up: mean 9 weeks	Study population		RR 0.78 (0.32 to 1.88)	199 (1 study)	⊕⊖⊖⊖ very low2,3,4
	102 per 1000	80 per 1000 (33 to 192)			
	Moderate				
	102 per 1000	80 per 1000 (33 to 192)			
Anxiety diagnosis Long Follow-up (25-103 weeks post-intervention) – ITT analysis MINI	Study population		RR 1 (0.56 to 1.78)	277 (1 study)	⊕⊖⊖⊖ very low1,2,3,4
	163 per 1000	163 per 1000 (91 to 290)			
	Moderate				
	163 per 1000	163 per 1000 (91 to 290)			

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Risk of bias as statistically significant group differences at baseline
2 Total number of events is less than 300 (a threshold rule-of-thumb)
3 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
4 Papers omit data

Anxiety: mother-infant relationship interventions versus treatment as usual or enhanced treatment as usual

There was low quality single study (N=98) evidence for a large effect of a mother–infant relationship intervention on anxiety symptomatology using an available case analysis (p=0.31). However, the imprecision of this effect estimate was very serious due to the small number of events and large 95% CI. In addition, when an ITT analysis approach was adopted there was no evidence for clinically or statistically significant benefits on anxiety symptomatology (p=0.86), or mean anxiety symptoms using an available case analysis at endpoint (p=0.44) or intermediate follow-up (p=0.15; Table 165).

Table 165: Summary of findings table for effects of mother–infant relationship interventions compared with treatment as usual or enhanced treatment as usual on anxiety outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Anxiety: Mother-infant relationship interventions versus TAU/Enhanced TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Anxiety symptomatology Post-treatment – ITT analysis STAI-S >40 Follow-up: mean 7 weeks	Study population 213 per 1000 Moderate 213 per 1000	200 per 1000 (100 to 403) 200 per 1000 (100 to 403)	RR 0.94 (0.47 to 1.89)	121 (1 study)	⊕⊕⊖⊖ low1,2	
Anxiety symptomatology Post-treatment – available case analysis STAI-S >40 Follow-up: mean 7 weeks	Study population 40 per 1000 Moderate 40 per 1000	8 per 1000 (0 to 169) 8 per 1000 (0 to 169)	RR 0.21 (0.01 to 4.23)	98 (1 study)	⊕⊕⊖⊖ low1,2	
Anxiety mean scores post-treatment – available case analysis STAI-S Follow-up: mean 7 weeks		The mean anxiety mean scores post-treatment – available case analysis in the intervention groups was 0.16 standard deviations lower (0.55 lower to 0.24 higher)		98 (1 study)	⊕⊕⊖⊖ low2,3	SMD -0.16 (-0.55 to 0.24)
Anxiety mean scores Intermediate follow-up (17-24 weeks post-intervention) – available case analysis STAI-S Follow-up: mean 25 weeks		The mean anxiety mean scores intermediate follow-up (17-24 weeks post-intervention) – available case analysis in the intervention groups was 0.3 standard deviations lower (0.7 lower to 0.11 higher)		96 (1 study)	⊕⊕⊖⊖ low2,3	SMD -0.3 (-0.7 to 0.11)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total number of events is less than 300 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
3 Total population size is less than 400 (a threshold rule-of-thumb)

Anxiety: music therapy during birth versus treatment as usual

There was low quality single study (N=141) evidence for a statistically significant large effect of music therapy during birth on anxiety mean symptoms immediately post-birth using an available case analysis approach (p <0.00001; Table 166). However, unfortunately, ITT (WCS) data cannot be extracted or computed for this outcome and meta-analysis was not possible. Moreover, the clinical significance and generalisability of effects on immediate post-birth anxiety to longer-term anxiety symptoms is unclear.

Table 166: Summary of findings table for effects of music therapy compared with treatment as usual on anxiety outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	 Corresponding risk Anxiety: Music therapy during birth versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Anxiety mean scores post-treatment – available case analysis Visual Analogue Scale (VAS) Anxiety Follow-up: mean 3 weeks		The mean anxiety mean scores post-treatment – available case analysis in the intervention groups was 2.16 standard deviations lower (2.58 to 1.74 lower)		141 (1 study)	⊕⊕⊖⊖ low[1]	SMD -2.16 (-2.58 to -1.74)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total population size is less than 400 (a threshold rule-of-thumb)

Anxiety: psychosomatic intervention versus treatment as usual

There was no evidence for a statistically or clinically significant effect of a psychosomatic intervention on mean anxiety symptoms (p=0.57; Table 167).

Table 167: Summary of findings table for effects of psychosomatic intervention compared with treatment as usual on anxiety outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	 Corresponding risk Anxiety: Psychosomatic intervention versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Anxiety mean scores post-treatment – available case analysis HADS – Anxiety Follow-up: mean 52 weeks		The mean anxiety mean scores post-treatment – available case analysis in the intervention groups was 0.17 standard deviations lower (0.76 lower to 0.42 higher)		44 (1 study)	⊕⊕⊖⊖ low1,2	SMD -0.17 (-0.76 to 0.42)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total population size is less than 400 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Anxiety: mindfulness training versus treatment as usual or enhanced treatment as usual

There was no evidence for statistically or clinically significant effects of mindfulness training on mean anxiety symptoms using either an ITT analysis (p=0.44) or available case analysis (p=0.95; Table 168).

Table 168: Summary of findings table for effects of mindfulness training compared with treatment as usual or enhanced treatment as usual on anxiety outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	 Corresponding risk Anxiety: Mindfulness training versus Enhanced TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Anxiety mean scores post-treatment – ITT analysis STAI-S Follow-up: mean 6 weeks		The mean anxiety mean scores post-treatment – ITT analysis in the intervention groups was 0.23 standard deviations higher (0.35 lower to 0.8 higher)		47 (1 study)	⊕⊕⊖⊖ low1,2	SMD 0.23 (-0.35 to 0.8)
Anxiety mean scores post-treatment – available case analysis STAI-S Follow-up: mean 10 weeks		The mean anxiety mean scores post-treatment – available case analysis in the intervention groups was 0.02 standard deviations lower (0.74 lower to 0.69 higher)		31 (1 study)	⊕⊖⊖⊖ very low1,2,3	SMD -0.02 (-0.74 to 0.69)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total population size is less than 400 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
3 Paper omits data

7.5.6 Clinical evidence for effects on adjustment disorder outcomes (by intervention)

Summary of findings can be found in the tables presented in this section. The full GRADE evidence profiles and associated forest plots can be found in Appendix 22 and Appendix 19, respectively.

Adjustment disorder: psychologically (CBT/IPT)-informed psychoeducation versus treatment as usual or enhanced treatment as usual

There was no evidence for a clinically or statistically significant effect of psychologically-informed psychoeducation on adjustment disorder diagnosis (p=0.77; Table 169).

Table 169: Summary of findings table for effects of psychologically (CBT/IPT)-informed psychoeducation compared with treatment as usual or enhanced treatment as usual on adjustment disorder outcomes

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Adjustment disorder: Psychologically (CBT/IPT)-informed psychoeducation versus TAU/Enhanced TAU				
Adjustment disorders diagnosis post-treatment – ITT analysis SADS Follow-up: mean 52 weeks	Study population		RR 0.9 (0.45 to 1.82)	199 (1 study)	⊕⊕⊖⊖ low1,2	
	143 per 1000	129 per 1000 (64 to 260)				
	Moderate					
	143 per 1000	129 per 1000 (64 to 260)				
Adjustment disorders diagnosis post-treatment – available case analysis SADS Follow-up: mean 52 weeks	Study population		RR 0.9 (0.45 to 1.82)	199 (1 study)	⊕⊕⊖⊖ low1,2	
	143 per 1000	129 per 1000 (64 to 260)				
	Moderate					
	143 per 1000	129 per 1000 (64 to 260)				

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total number of events is less than 300 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

7.5.7 Clinical evidence for effects on PTSD outcomes (by intervention)

Summary of findings can be found in the tables presented in this section. The full GRADE evidence profiles and associated forest plots can be found in Appendix 22 and Appendix 19, respectively.

PTSD: post-miscarriage self-help versus treatment as usual

There was low quality, single study (N=78) evidence for moderate to large effects of post-miscarriage self-help on PTSD symptomatology (analysed using ITT [p=0.02] or available case [p=0.004] approaches) and large effects on mean PTSD symptoms (p=0.0004; Table 170).

Table 170: Summary of findings table for effects of post-miscarriage self-help compared with treatment as usual on PTSD outcomes

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	PTSD: post-miscarriage self-help versus TAU				
PTSD symptomatology Post-treatment – ITT analysis IES: Treatment non-response (reliable change index) Follow-up: mean 5 weeks	Study population		RR 0.59 (0.38 to 0.94)	78 (1 study)	⊕⊕⊖⊖ low1	
	636 per 1000	375 per 1000 (242 to 598)				
	Moderate					
	636 per 1000	375 per 1000 (242 to 598)				

PTSD symptomatology Post-treatment – available case analysis IES: Treatment non-response (reliable change index) Follow-up: mean 5 weeks	Study population		RR 0.32 (0.14 to 0.7)	59 (1 study)	⊕⊕⊖⊖ low1	
	577 per 1000	185 per 1000 (81 to 404)				
	Moderate					
	577 per 1000	185 per 1000 (81 to 404)				
PTSD mean scores post-treatment – ITT analysis IES: Traumatic stress Follow-up: mean 5 weeks		The mean PTSD mean scores post-treatment – ITT analysis in the intervention groups was 0.84 standard deviations lower (1.31 to 0.37 lower)		78 (1 study)	⊕⊕⊖⊖ low2	SMD -0.84 (-1.31 to -0.37)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total number of events is less than 300 (a threshold rule-of-thumb)
2 Total population size is less than 400 (a threshold rule-of-thumb)

PTSD: post-traumatic birth counselling versus treatment as usual

There was no evidence for statistically or clinically significant benefits of post-traumatic birth counselling on PTSD diagnosis (p=0.10) and no evidence for a clinically significant effect (despite meeting statistical significance criteria as p=0.04) on mean PTSD symptoms (Table 171).

Table 171: Summary of findings table for effects of post-traumatic counselling compared with treatment as usual on PTSD outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk PTSD: post-traumatic birth counselling versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
PTSD diagnosis post-treatment – ITT analysis Mini- PTSD Diagnosis Interview Follow-up: mean 13 weeks	Study population		RR 0.35 (0.1 to 1.23)	103 (1 study)	⊕⊕⊖⊖ low1,2	
	170 per 1000	59 per 1000 (17 to 209)				
	Moderate					
	170 per 1000	59 per 1000 (17 to 209)				
PTSD diagnosis post-treatment – available case analysis Mini- PTSD Diagnosis Interview Follow-up: mean 13 weeks	Study population		RR 0.35 (0.1 to 1.23)	103 (1 study)	⊕⊕⊖⊖ low1,2	
	170 per 1000	59 per 1000 (17 to 209)				
	Moderate					
	170 per 1000	59 per 1000 (17 to 209)				
PTSD mean scores post-treatment – ITT analysis Mini- PTSD Diagnosis Interview: 'Trauma symptoms', rating scale unclear Follow-up: mean 13 weeks		The mean PTSDPTSD mean scores post-treatment – ITT analysis in the intervention groups was 0.41 standard deviations lower (0.81 to 0.02 lower)		103 (1 study)	⊕⊕⊖⊖ low3	SMD -0.41 (-0.81 to -0.02)
PTSD mean scores post-treatment – available case analysis Mini- PTSD Diagnosis Interview: 'Trauma symptoms', rating scale unclear		The mean PTSD mean scores post-treatment – available case analysis in the intervention groups was 0.41 standard deviations lower (0.81 to 0.02 lower)		103 (1 study)	⊕⊕⊖⊖ low3	SMD -0.41 (-0.81 to -0.02)

Follow-up: mean 13 weeks					

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total number of events is less than 300 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
3 Total population size is less than 400 (a threshold rule-of-thumb)

PTSD: psychologically (CBT/IPT)-informed psychoeducation versus treatment as usual or enhanced treatment as usual

There was inconsistent evidence for benefits associated with psychoeducation for PTSD outcomes, with the ITT analysis of PTSD symptomatology suggestive of moderate benefits of psychoeducation (p=0.63), the available case analysis suggestive of large harms associated with psychoeducation for PTSD symptomatology (p=0.56), and two studies (N=96) providing evidence for small benefits of psychoeducation on continuous measures of PTSD symptoms (p=0.05). However, there was no evidence for statistically significant benefits for any of the outcome measures and the very low quality of evidence due to risk of bias concerns (unclear blinding of outcome assessment), very serious imprecision (due to small event rates/sample size and large 95% CIs) and selective outcome reporting prohibits any clear conclusions being drawn from this evidence (Table 172).

Table 172: Summary of findings table for effects of psychologically (CBT/IPT)-informed psychoeducation compared with treatment as usual on PTSD outcomes

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	PTSD: Psychologically (CBT/IPT)-informed psychoeducation versus TAU/Enhanced TAU				
PTSD diagnosis post-treatment – ITT analysis	Study population		RR 0.74 (0.22 to 2.47)	54 (1 study)	⊕⊖⊖⊖ very low1,2,3,4	
	192 per 1000	142 per 1000 (42 to 475)				

LIFE Follow-up: mean 13 weeks	Moderate 192 per 1000	142 per 1000 (42 to 474)			
PTSD diagnosis post-treatment – available case analysis LIFE Follow-up: mean 13 weeks	Study population 0 per 1000 Moderate 0 per 1000	0 per 1000 (0 to 0) 0 per 1000 (0 to 0)	RR 2.54 (0.11 to 59.23)	46 (1 study)	⨁⊖⊖⊖ very low1,2,3,4	
PTSD mean scores post-treatment – available case analysis Davidson Trauma Scale or LIFE: Psychiatric Status Ratings mean PTSD score Follow-up: 6-13 weeks		The mean PTSD mean scores post-treatment – available case analysis in the intervention groups was 0.4 standard deviations lower (0.81 lower to 0 higher)		96 (2 studies)	⨁⊖⊖⊖ very low4,5	SMD -0.4 (-0.81 to 0)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Risk of bias due to unclear blinding of outcome assessment
2 Total number of events is less than 300 (a threshold rule-of-thumb)
3 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
4 Papers omit data
5 Total population size is less than 400 (a threshold rule-of-thumb)

PTSD: mother-infant relationship interventions versus treatment as usual or enhanced treatment as usual

There was no evidence for clinically or statistically significant benefits or harms associated with mother–infant relationship interventions for PTSD symptomatology at endpoint when an ITT analysis approach was adopted (p=0.52) or at intermediate follow-up using either data analysis method (p=0.57-0.95) or for PTSD mean symptoms at endpoint (p=0.61) or intermediate follow-up (p=0.21). There was low quality single study (N=98) evidence for moderate harms associated with a mother–infant relationship intervention on PTSD symptomatology when an available case analysis was used (p=0.54). However, very serious imprecision of this effect estimate prohibits any clear conclusions being drawn from this data (Table 173).

Table 173: Summary of findings table for effects of mother–infant relationship interventions compared with treatment as usual on PTSD outcomes

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	PTSD: Mother-infant relationship interventions versus TAU/Enhanced TAU				
PTSD symptomatology Post-treatment – ITT analysis PPQ: Scores in clinical range (no further detail) Follow-up: mean 7 weeks	Study population		RR 1.18 (0.71 to 1.94)	121 (1 study)	⊕⊕⊖⊖ low1,2	
	311 per 1000	368 per 1000 (221 to 604)				
	Moderate					
	312 per 1000	368 per 1000 (222 to 605)				
PTSD symptomatology Post-treatment – available case analysis PPQ: Scores in clinical range (no further detail) Follow-up: mean 7 weeks	Study population		RR 1.3 (0.56 to 3.02)	98 (1 study)	⊕⊕⊖⊖ low1,2	
	160 per 1000	208 per 1000 (90 to 483)				
	Moderate					
	160 per 1000	208 per 1000 (90 to 483)				
PTSD mean scores post-treatment – available case analysis PPQ Follow-up: mean 7 weeks		The mean PTSD mean scores post-treatment – available case analysis in the intervention groups was 0.1 standard deviations lower (0.5 lower to 0.29 higher)		98 (1 study)	⊕⊕⊖⊖ low2,3	SMD -0.1 (-0.5 to 0.29)
PTSD symptomatology Intermediate follow-up (17-24 weeks post-intervention) – ITT analysis PPQ: Scores in clinical range (no further detail) Follow-up: mean 25 weeks	Study population		RR 1.02 (0.63 to 1.63)	121 (1 study)	⊕⊕⊖⊖ low1,2	
	361 per 1000	368 per 1000 (227 to 588)				
	Moderate					
	361 per 1000	368 per 1000 (227 to 588)				

PTSD symptomatology Intermediate follow-up (17-24 weeks post-intervention) – available case analysis PPQ: Scores in clinical range (no further detail) Follow-up: mean 25 weeks	Study population 220 per 1000	174 per 1000 (77 to 394)	RR 0.79 (0.35 to 1.79)	96 (1 study)	⊕⊕⊖⊖ low1,2	
	Moderate 220 per 1000	174 per 1000 (77 to 394)				
PTSD mean scores Intermediate follow-up (17-24 weeks post-intervention) – available case analysis PPQ Follow-up: mean 25 weeks		The mean PTSD mean scores intermediate follow-up (17-24 weeks post-intervention) – available case analysis in the intervention groups was 0.25 standard deviations lower (0.66 lower to 0.15 higher)		96 (1 study)	⊕⊕⊖⊖ low2,3	SMD -0.25 (-0.66 to 0.15)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total number of events is less than 300 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
3 Total population size is less than 400 (a threshold rule-of-thumb)

7.5.8 Clinical evidence for effects on OCD outcomes (by intervention)

Summary of findings can be found in the tables presented in this section. The full GRADE evidence profiles and associated forest plots can be found in Appendix 22 and Appendix 19, respectively.

OCD: psychologically (CBT/IPT)-informed psychoeducation versus treatment as usual or enhanced treatment as usual

There was very low quality single study (N=58) evidence for delayed but statistically significant moderate to large effects of psychoeducation on mean OCD symptoms at intermediate and long-term follow-ups (total scores [p=0.01-0.02] and obsessions [p=0.02-0.03] and compulsions [p=0.02] subscales), with statistically and clinically non-significant effects at endpoint (p=0.12-0.24; Table 174).

Table 174: Summary of findings table for effects of psychologically (CBT/IPT)-informed psychoeducation interventions compared with treatment as usual or enhanced treatment as usual on OCD outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk OCD: Psychologically (CBT/IPT)-informed psychoeducation versus TAU/Enhanced TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
OCD mean scores post-treatment – available case analysis Yale–Brown Obsessive Compulsive Scale (YBOCS) Follow-up: mean 4 weeks		The mean OCD mean scores post-treatment – available case analysis in the intervention groups was 0.41 standard deviations lower (0.94 lower to 0.11 higher)		58 (1 study)	⊕⊖⊖⊖ very low1,2,3	SMD -0.41 (-0.94 to 0.11)
Obsessions mean scores post-treatment – available case analysis YBOCS: Obsessions Follow-up: mean 4 weeks		The mean obsessions mean scores post-treatment – available case analysis in the intervention groups was 0.39 standard deviations lower (0.92 lower to 0.13 higher)		58 (1 study)	⊕⊖⊖⊖ very low1,2,3	SMD -0.39 (-0.92 to 0.13)
Compulsions mean scores post-treatment – available case analysis YBOCS: Compulsions Follow-up: mean 4 weeks		The mean compulsions mean scores post-treatment – available case analysis in the intervention groups was 0.31 standard deviations lower (0.83 lower to 0.21 higher)		58 (1 study)	⊕⊖⊖⊖ very low1,2,3	SMD -0.31 (-0.83 to 0.21)
OCD mean scores Intermediate follow-up (17-24 weeks post-intervention) – available case analysis YBOCS Follow-up: mean 19 weeks		The mean OCD mean scores intermediate follow-up (17-24 weeks post-intervention) – available case analysis in the intervention groups was 0.71 standard deviations lower (1.29 to 0.12 lower)		50 (1 study)	⊕⊖⊖⊖ very low1,3	SMD -0.71 (-1.29 to -0.12)

Obsessions mean scores Intermediate follow-up (17-24 weeks post-intervention) – available case analysis YBOCS: Obsessions Follow-up: mean 19 weeks	The mean obsessions mean scores intermediate follow-up (17-24 weeks post-intervention) – available case analysis in the intervention groups was 0.65 standard deviations lower (1.24 to 0.07 lower)	50 (1 study)	⊕⊖⊖⊖ very low1,3	SMD -0.65 (-1.24 to -0.07)
Compulsions mean scores Intermediate follow-up (17-24 weeks post-intervention) – available case analysis YBOCS: Compulsions Follow-up: mean 19 weeks	The mean compulsions mean scores intermediate follow-up (17-24 weeks post-intervention) – available case analysis in the intervention groups was 0.7 standard deviations lower (1.29 to 0.11 lower)	50 (1 study)	⊕⊖⊖⊖ very low1,3	SMD -0.7 (-1.29 to -0.11)
OCD mean scores Long follow-up (25-103 weeks post-intervention) – available case analysis YBOCS Follow-up: mean 32 weeks	The mean OCD mean scores long follow-up (25-103 weeks post-intervention) – available case analysis in the intervention groups was 0.76 standard deviations lower (1.35 to 0.17 lower)	49 (1 study)	⊕⊖⊖⊖ very low1,3	SMD -0.76 (-1.35 to -0.17)
Obsessions mean scores Long follow-up (25-103 weeks post-intervention) – available case analysis YBOCS: Obsessions Follow-up: mean 32 weeks	The mean obsessions mean scores long follow-up (25-103 weeks post-intervention) – available case analysis in the intervention groups was 0.73 standard deviations lower (1.32 to 0.14 lower)	49 (1 study)	⊕⊖⊖⊖ very low1,3	SMD -0.73 (-1.32 to -0.14)
Compulsions mean scores Long follow-up (25-103 weeks post-intervention) – available case analysis YBOCS: Compulsions Follow-up: mean 32 weeks	The mean compulsions mean scores long follow-up (25-103 weeks post-intervention) – available case analysis in the intervention groups was 0.72 standard deviations lower (1.31 to 0.13 lower)	49 (1 study)	⊕⊕⊖⊖ low1	SMD -0.72 (-1.31 to -0.13)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total population size is less than 400 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
3 Paper omits data

7.5.9 Clinical evidence for effects on fear of childbirth outcomes (by intervention)

Summary of findings can be found in the tables presented in this section. The full GRADE evidence profiles and associated forest plots can be found in Appendix 22 and Appendix 19, respectively.

Fear of childbirth: pre-delivery discussion/psychoeducation versus treatment as usual

There was no evidence for clinically or statistically significant benefits of pre-delivery discussion/psychoeducation on mode of delivery (elective caesarean [p=0.76]; choosing vaginal delivery [p=0.69]; vaginal delivery [p=0.21]) or for pre-delivery fear of, or preparedness for, childbirth (p=0.13-0.53) or satisfaction with childbirth (p=0.14). There was moderate to very low quality, single study (N=176-371) evidence for small but statistically significant effects on continuous measures of feeling safe during childbirth (p=0.01), experience of fear during childbirth (p=0.001), and maternal attitude to motherhood (p=0.02). However, these benefits were not appreciable and may not be clinically meaningful (Table 175).

Table 175: Summary of findings table for effects of pre-delivery discussion/psychoeducation compared with treatment as usual on fear of childbirth outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Fear of childbirth: Pre-delivery discussion/psychoeducation versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Elective caesarean Post-treatment – ITT analysis Mode of delivery: Number of women delivering via elective caesarean or caesarean for psychosocial reasons Follow-up: 0-16 weeks	Study population		RR 0.93 (0.57 to 1.51)	461 (2 studies)	⊕⊕⊖⊖ low1,2	
	136 per 1000	127 per 1000 (78 to 206)				
	Moderate					
	152 per 1000	141 per 1000 (87 to 230)				
Choosing vaginal delivery Post-treatment – ITT analysis Delivery preference: Number of women choosing vaginal delivery Follow-up: mean 16 weeks	Study population		RR 1.05 (0.84 to 1.3)	90 (1 study)	⊕⊖⊖⊖ very low1,2,3	
	761 per 1000	799 per 1000 (639 to 989)				
	Moderate					
	761 per 1000	799 per 1000 (639 to 989)				
Vaginal delivery Post-treatment – ITT analysis Mode of delivery: Spontaneous vaginal delivery/vaginal delivery Follow-up: 0-16 weeks	Study population		RR 1.2 (0.9 to 1.59)	462 (2 studies)	⊕⊖⊖⊖ very low1,2,4	
	491 per 1000	590 per 1000 (442 to 781)				
	Moderate					
	525 per 1000	630 per 1000 (472 to 835)				

Fear of pain in labour mean score Mid-treatment (36 weeks gestation) – ITT analysis Pregnancy Anxiety Scale: Fear of pain in labour Follow-up: mean 12 weeks	The mean fear of pain in labour mean score mid-treatment (36 weeks gestation) – ITT analysis in the intervention groups was 0.09 standard deviations lower (0.39 lower to 0.2 higher)	176 (1 study)	⊕⊖⊖⊖ very low3,5	SMD -0.09 (-0.39 to 0.2)
Fear of obstetrician's unfriendly behaviour mean scores Mid-treatment (36 weeks gestation) – ITT analysis Pregnancy Anxiety Scale: Fear of obstetrician's unfriendly behaviour Follow-up: mean 12 weeks	The mean fear of obstetrician's unfriendly behaviour mean scores mid-treatment (36 weeks gestation) – ITT analysis in the intervention groups was 0.23 standard deviations lower (0.53 lower to 0.07 higher)	176 (1 study)	⊕⊖⊖⊖ very low2,3,5	SMD -0.23 (-0.53 to 0.07)
Preparedness for childbirth mean scores Mid-treatment (36 weeks gestation) – available case analysis Preparedness for childbirth (study-specific scale) Follow-up: mean 8 weeks	The mean preparedness for childbirth mean scores mid-treatment (36 weeks gestation) – available case analysis in the intervention groups was 0.19 standard deviations higher (0.07 lower to 0.44 higher)	254 (1 study)	⊕⊕⊕⊖ moderate5	SMD 0.19 (-0.07 to 0.44)
Satisfaction with childbirth mean scores post-treatment – ITT analysis Study-specific scale: Satisfaction with childbirth Follow-up: mean 29 weeks	The mean satisfaction with childbirth mean scores post-treatment – ITT analysis in the intervention groups was 0.22 standard deviations lower (0.52 lower to 0.08 higher)	176 (1 study)	⊕⊖⊖⊖ very low2,3,5	SMD -0.22 (-0.52 to 0.08)
Feeling safe during childbirth mean scores post-treatment – ITT analysis	The mean feeling safe during childbirth mean scores post-treatment – ITT analysis in the intervention groups was 0.39 standard deviations	176 (1 study)	⊕⊖⊖⊖ very low3,5	SMD -0.39 (-0.69 to -0.09)

Satisfaction with childbirth: Feeling safe (study-specific scale) Follow-up: mean 29 weeks	lower (0.69 to 0.09 lower)			
Experience of fear during childbirth mean scores post-treatment – ITT analysis Wilma Delivery Experience Questionnaire (W-DEQ-B) Follow-up: mean 13 weeks	The mean experience of fear during childbirth mean scores post-treatment – ITT analysis in the intervention groups was 0.35 standard deviations lower (0.57 to 0.14 lower)	371 (1 study)	⊕⊕⊕⊖ moderate5	SMD -0.35 (-0.57 to -0.14)
Maternal attitude to motherhood mean scores post-treatment – available case analysis Motherhood and parenting (based on Kumar et al., 1984) Follow-up: mean 25 weeks	The mean maternal attitude to motherhood mean scores post-treatment – available case analysis in the intervention groups was 0.3 standard deviations higher (0.04 to 0.56 higher)	252 (1 study)	⊕⊕⊕⊖ moderate5	SMD 0.3 (0.04 to 0.56)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total number of events is less than 300 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
3 Paper omits data
4 There was evidence of moderate heterogeneity between effect sizes
5 Total population size is less than 400 (a threshold rule-of-thumb)

7.5.10 Clinical evidence for effects on eating disorder outcomes (by intervention)

Summary of findings can be found in the tables presented in this section. The full GRADE evidence profiles and associated forest plots can be found in Appendix 22 and Appendix 19, respectively.

Eating disorders: mother-infant relationship interventions (and facilitated self-help) versus listening visits (and facilitated self-help)

There was no evidence for statistically or clinically significant benefits of mother–infant relationship interventions compared with listening visits on eating disorder diagnosis (p=0.81-0.92; Table 176). However, it is important to note that participants in both active intervention arms received facilitated self-help aimed at their eating disorder.

Table 176: Summary of findings table for effects of mother–infant relationship intervention (and facilitated self-help) compared with listening visits (and facilitated self-help) on eating disorder outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Eating disorder: Mother-infant relationship interventions (and facilitated self-help) versus listening visits (and facilitated self-help)	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Eating disorder diagnosis post-treatment – ITT analysis Psychiatric interview: DSM-IV Eating Disorder Follow-up: mean 35 weeks	Study population		RR 1.08 (0.58 to 1.99)	80 (1 study)	⊕⊖⊖⊖ very low1,2,3	
	325 per 1000	351 per 1000 (188 to 647)				
	Moderate					
	325 per 1000	351 per 1000 (188 to 647)				
Eating disorder diagnosis post-treatment – available case analysis Psychiatric interview: DSM-IV Eating Disorder	Study population		RR 0.97 (0.49 to 1.91)	76 (1 study)	⊕⊖⊖⊖ very low1,2,3	
	308 per 1000	298 per 1000 (151 to 588)				
	Moderate					
	308 per 1000	299 per 1000 (151 to 588)				

Follow-up: mean 35 weeks	

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total number of events is less than 300 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
3 Paper omits data

7.5.11 Clinical evidence for effects on general mental health outcomes (by intervention)

Summary of findings can be found in the tables presented in this section. The full GRADE evidence profiles and associated forest plots can be found in Appendix 22 and Appendix 19, respectively.

General mental health outcomes: structured psychological interventions (CBT or IPT) versus treatment as usual or enhanced treatment as usual

There was low to very low quality evidence from up to two studies (N=305) for moderate to large benefits of structured psychological interventions (CBT or IPT) on general mental health outcomes at endpoint (p=0.0004-0.08), and at short-term (p=0.0007) and intermediate (p=0.06) follow-ups. There was also evidence for a statistically significant, but not clinically significant, effect of CBT on reducing the risk of self-harm (p=0.009) (Table 177).

Table 177: Summary of findings table for effects of structured psychological interventions (CBT or IPT) compared with treatment as usual or enhanced treatment as usual on general mental health outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk General mental health: Structured psychological interventions (CBT or IPT) versus TAU/Enhanced TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
General mental health mean scores post-treatment – ITT analysis BSI: Global severity index (Mental health) Follow-up: mean 15 weeks		The mean general mental health mean scores post-treatment – ITT analysis in the intervention groups was 0.76 standard deviations lower (1.19 to 0.34 lower)		93 (1 study)	⊕⊕⊖⊖ low1	SMD -0.76 (-1.19 to -0.34)
General mental health (higher better) mean scores post-treatment – available case analysis SF-12 Mental component summary Follow-up: 15-26 weeks		The mean general mental health (higher better) mean scores post-treatment – available case analysis in the intervention groups was 0.68 standard deviations higher (0.08 lower to 1.44 higher)		305 (2 studies)	⊕⊖⊖⊖ very low1,2,3,4	SMD 0.68 (-0.08 to 1.44)
Risk of self-harm mean scores post-treatment – available case analysis CORE-OM: Risk of self-harm Follow-up: mean 26 weeks		The mean risk of self-harm mean scores post-treatment – available case analysis in the intervention groups was 0.31 standard deviations lower (0.55 to 0.08 lower)		283 (1 study)	⊕⊕⊖⊖ low1,4	SMD -0.31 (-0.55 to -0.08)
General mental health mean scores Short follow-up (9-16 weeks post-intervention) – ITT analysis BSI: Global severity		The mean general mental health mean scores short follow-up (9-16 weeks post-intervention) – ITT analysis in the intervention groups was 0.73 standard deviations		93 (1 study)	⊕⊕⊖⊖ low1	SMD -0.73 (-1.15 to -0.31)

index (Mental health) Follow-up: mean 28 weeks	lower (1.15 to 0.31 lower)			
General mental health mean scores Intermediate follow-up (17-24 weeks post-intervention) – available case analysis SF-12 Mental component summary Follow-up: mean 33 weeks	The mean general mental health mean scores intermediate follow-up (17-24 weeks post-intervention) – available case analysis in the intervention groups was 0.78 standard deviations higher (0.03 lower to 1.59 higher)	26 (1 study)	⊕⊖⊖⊖ very low1,3,5	SMD 0.78 (-0.03 to 1.59)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total population size is less than 400 (a threshold rule-of-thumb)
2 There was evidence of substantial heterogeneity between effect sizes
3 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
4 Papers omit data
5 Risk of bias due to statistically significant group differences at baseline

General mental health outcomes: IPT versus support group

There was no evidence for clinically or statistically significant benefits of IPT relative to a support group on anger mean scores (p=0.77; Table 178).

Table 178: Summary of findings table for effects of IPT compared with support group on general mental health outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk General mental health: IPT versus support group	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Anger Post-treatment (mean score at endpoint or first measurement) – available case analysis State Anger Inventory (STAXI) Follow-up: mean 12 weeks		The mean anger post-treatment (mean score at endpoint or first measurement) – available case analysis in the intervention groups was 0.09 standard deviations lower (0.68 lower to 0.5 higher)		44 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	SMD -0.09 (-0.68 to 0.5)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Risk of bias due to statistically significant group differences at baseline
[2] Total population size is less than 400 (a threshold rule-of-thumb)
[3] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

General mental health outcomes: post-miscarriage self-help versus treatment as usual

There was single study (N=78) evidence for moderate to large effects of post-miscarriage self-help on global mental health severity (treatment non-response [p=0.02-0.06] and mean scores [p=0.005]) (Table 179).

Table 179: Summary of findings table for effects of post-miscarriage self-help compared with treatment as usual on general mental health outcomes

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	General mental health: post-miscarriage self-help versus TAU				
General mental health Post-treatment (treatment non-response/symptomatology at endpoint or first measurement) – ITT analysis BSI: Global severity index (Treatment non-response: reliable change index) Follow-up: mean 5 weeks	Study population		RR 0.7 (0.48 to 1.02)	78 (1 study)	⊕⊕⊖⊖ low1,2	
	697 per 1000	488 per 1000 (335 to 711)				
	Moderate					
	697 per 1000	488 per 1000 (335 to 711)				
General mental health Post-treatment (treatment non-response/symptomatology at endpoint or first measurement) – available case analysis BSI: Global severity index (Treatment non-response: reliable change index) Follow-up: mean 5 weeks	Study population		RR 0.49 (0.27 to 0.9)	59 (1 study)	⊕⊕⊖⊖ low1	
	615 per 1000	302 per 1000 (166 to 554)				
	Moderate					
	615 per 1000	301 per 1000 (166 to 553)				
General mental health Post-treatment (mean mental health symptoms at endpoint or first measurement) – ITT analysis BSI: Global severity index (Mental health) Follow-up: mean 5 weeks		The mean general mental health post-treatment (mean mental health symptoms at endpoint or first measurement) – ITT analysis in the intervention groups was 0.67 standard deviations lower (1.13 to 0.21 lower)		78 (1 study)	⊕⊕⊖⊖ low3	SMD -0.67 (-1.13 to -0.21)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.

Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total number of events is less than 300 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
3 Total population size is less than 400 (a threshold rule-of-thumb)

General mental health outcomes: listening visits versus treatment as usual

There was single study (N=271-276) evidence for small benefits of listening visits on general mental health (p=0.0006) and risk of self-harm (p=0.01) mean scores (Table 180). However, these effects are too small to meet criteria for appreciable benefits and are unlikely to be clinically meaningful.

Table 180: Summary of findings table for effects of listening visits compared with treatment as usual on general mental health outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk General mental health: Listening visits versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
General mental health (higher better) Post-treatment (mean mental health symptoms at endpoint or first measurement) – available case analysis SF-12 Mental component summary Follow-up: mean 26 weeks		The mean general mental health (higher better) post-treatment (mean mental health symptoms at endpoint or first measurement) – available case analysis in the intervention groups was 0.42 standard deviations higher (0.18 to 0.66 higher)		271 (1 study)	⊕⊕⊖⊖ low[1,2]	SMD 0.42 (0.18 to 0.66)
Risk of self-harm Post-treatment (mean score at endpoint or first measurement) – available case analysis		The mean risk of self-harm post-treatment (mean score at endpoint or first measurement) – available case analysis in the intervention groups was		276 (1 study)	⊕⊕⊖⊖ low[1,2]	SMD -0.31 (-0.55 to -0.07)

CORE-OM: Risk of self-harm Follow-up: mean 26 weeks	0.31 standard deviations lower (0.55 to 0.07 lower)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total population size is less than 400 (a threshold rule-of-thumb)
[2] Paper omits data

General mental health outcomes: post-miscarriage counselling versus treatment as usual

There was no evidence for clinically or statistically significant effects of post-miscarriage counselling on feelings of self-blame at post-treatment (p=0.55) or intermediate follow-up (p=0.91) (Table 181).

Table 181: Summary of findings table for effects of post-miscarriage counselling compared with treatment as usual on general mental health outcomes

Outcomes	**Illustrative comparative risks* (95% CI)** **Assumed risk** **Control**	**Corresponding risk** **General mental health: post-miscarriage counselling versus TAU**	**Relative effect (95% CI)**	**No. of Participants (studies)**	**Quality of the evidence (GRADE)**	**Comments**
Self-blame Post-treatment (mean score at endpoint or first measurement) – available case analysis Study-specific measure: Self-blame Follow-up: mean 2 weeks		The mean self-blame post-treatment (mean score at endpoint or first measurement) – available case analysis in the intervention groups was 0.15 standard deviations higher (0.34 lower to 0.63 higher)		66 (1 study)	⊕⊕⊖⊖ low[1,2]	SMD 0.15 (-0.34 to 0.63)
Self-blame Intermediate		The mean self-blame intermediate follow-up		66 (1 study)	⊕⊕⊖⊖ low[1,2]	SMD 0.03 (-0.45 to 0.51)

follow-up (mean score at 17-24 week follow-up) – available case analysis Study-specific measure: Self-blame Follow-up: mean 17 weeks	(mean score at 17-24 week follow-up) – available case analysis in the intervention groups was 0.03 standard deviations higher (0.45 lower to 0.51 higher)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total population size is less than 400 (a threshold rule-of-thumb)
[2] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

General mental health outcomes: post-traumatic birth counselling versus treatment as usual

There was low quality, single study (N=103) evidence for large harms associated with post-traumatic birth counselling (p <0.00001) with mean scores on a study-specific measure of feelings of self-blame favouring treatment as usual (Table 182).

Table 182: Summary of findings table for effects of post-traumatic birth counselling compared with treatment as usual on general mental health outcomes

Outcomes	**Illustrative comparative risks* (95% CI)** **Assumed risk** **Control**	**Corresponding risk** **General mental health: post-traumatic birth counselling versus TAU**	**Relative effect (95% CI)**	**No. of Participants (studies)**	**Quality of the evidence (GRADE)**	**Comments**
Self-blame Post-treatment (feelings of self-blame at endpoint or first measurement) – ITT analysis Study-specific		The mean self-blame post-treatment (feelings of self-blame at endpoint or first measurement) – ITT analysis in the intervention groups was		103 (1 study)	⊕⊕⊖⊖ low1	SMD 2.37 (1.86 to 2.88)

measure: Self-blame Follow-up: mean 13 weeks	2.37 standard deviations higher (1.86 to 2.88 higher)			
Self-blame Post-treatment (feelings of self-blame at endpoint or first measurement) – available case analysis Study-specific measure: Self-blame Follow-up: mean 13 weeks	The mean self-blame post-treatment (feelings of self-blame at endpoint or first measurement) – available case analysis in the intervention groups was 2.37 standard deviations higher (1.86 to 2.88 higher)	103 (1 study)	⊕⊕⊖⊖ low1	SMD 2.37 (1.86 to 2.88)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total population size is less than 400 (a threshold rule-of-thumb)

General mental health outcomes: psychologically (CBT/IPT)-informed psychoeducation versus treatment as usual or enhanced treatment as usual

There was no evidence for clinically significant benefits (or harms) of psychoeducation on diagnosis of any psychopathology (p=0.90) or on general mental health mean scores at post-treatment (p=0.001) or short-term follow-up (p=0.27) (Table 183).

Table 183: Summary of findings table for effects of psychologically (CBT/IPT)-informed psychoeducation compared with treatment as usual or enhanced treatment as usual on general mental health outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk General mental health: Psychologically (CBT/IPT)-informed psychoeducation versus TAU/Enhanced TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Any psychopathology diagnosis post-treatment – ITT analysis SADS: Any psychopathology Follow-up: mean 52 weeks	Study population 367 per 1000 Moderate 367 per 1000	375 per 1000 (261 to 540) 374 per 1000 (261 to 539)	RR 1.02 (0.71 to 1.47)	199 (1 study)	⊕⊖⊖⊖ very low1,2,3	
Any psychopathology diagnosis post-treatment – available case analysis SADS: Any psychopathology Follow-up: mean 52 weeks	Study population 367 per 1000 Moderate 367 per 1000	375 per 1000 (261 to 540) 374 per 1000 (261 to 539)	RR 1.02 (0.71 to 1.47)	199 (1 study)	⊕⊖⊖⊖ very low1,2,3	
General mental health mean scores post-treatment – ITT analysis GHQ Follow-up: mean 6 weeks		The mean general mental health mean scores post-treatment – ITT analysis in the intervention groups was 0.48 standard deviations lower (0.76 to 0.19 lower)		194 (1 study)	⊕⊕⊖⊖ low4	SMD -0.48 (-0.76 to -0.19)
General mental health mean scores Short follow-up (9-16 weeks post-intervention) – ITT analysis GHQ Follow-up: mean 13 weeks		The mean general mental health mean scores short follow-up (9-16 weeks post-intervention) – ITT analysis in the intervention groups was 0.16 standard		194 (1 study)	⊕⊕⊖⊖ low4	SMD -0.16 (-0.44 to 0.12)

		deviations lower (0.44 lower to 0.12 higher)				

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total number of events is less than 300 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
3 Paper omits data
4 Total population size is less than 400 (a threshold rule-of-thumb)

General mental health outcomes: home visits versus treatment as usual or enhanced treatment as usual

There was no evidence of clinically or statistically significant benefits of home visits on general mental health symptomatology (p=0.47-0.79) or on alcohol or drug use (p=0.22-0.34) (Table 184).

Table 184: Summary of findings table for effects of home visits compared with treatment as usual or enhanced treatment as usual on general mental health outcomes

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	General mental health: Home visits versus TAU/Enhanced TAU				
General mental health symptomatology/treatment non-response Post-treatment – ITT analysis Mental Health Index – 5 <67 Follow-up: mean 104 weeks	Study population		RR 0.93 (0.77 to 1.13)	364 (1 study)	⊕⊕⊖⊖ low[1,2]	
	546 per 1000	508 per 1000 (420 to 617)				
	Moderate					
	546 per 1000	508 per 1000 (420 to 617)				
	Study population					

General mental health symptomatology/treatment non-response Post-treatment – available case analysis Mental Health Index <67 Follow-up: mean 104 weeks	317 per 1000	301 per 1000 (209 to 438)	RR 0.95 (0.66 to 1.38)	249 (1 study)	⊕⊖⊖⊖ very low[1,2,3]
	Moderate				
	317 per 1000	301 per 1000 (209 to 437)			
Alcohol or drug use symptomatology Post-treatment – ITT analysis CAGE Questionnaire: Alcohol or drug use Follow-up: mean 104 weeks	Study population		RR 0.88 (0.73 to 1.08)	364 (1 study)	⊕⊖⊖⊖ very low[1,2,3]
	557 per 1000	490 per 1000 (406 to 601)			
	Moderate				
	557 per 1000	490 per 1000 (407 to 602)			
Alcohol or drug use symptomatology Post-treatment – available case analysis CAGE Questionnaire: Alcohol or drug use Follow-up: mean 104 weeks	Study population		RR 0.83 (0.57 to 1.21)	249 (1 study)	⊕⊖⊖⊖ very low[1,2,3]
	333 per 1000	277 per 1000 (190 to 403)			
	Moderate				
	333 per 1000	276 per 1000 (190 to 403)			

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Risk of bias due to statistically significant group differences at baseline
[2] Total number of events is less than 300 (a threshold rule-of-thumb)
[3] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

General mental health outcomes: mother-infant relationship interventions versus treatment as usual or enhanced treatment as usual

There was no evidence for clinically or statistically significant effects of mother–infant relationship interventions on general mental health treatment non-response (p=0.42-0.50) or global severity mean scores (p=0.29) (Table 185).

Table 185: Summary of findings table for effects of mother–infant relationship interventions compared with treatment as usual or enhanced treatment as usual on general mental health outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk General mental health: Mother-infant relationship interventions versus TAU/Enhanced TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
General mental health treatment non-response Post-treatment – ITT analysis Symptom Checklist-90 (SCL-90): Global Severity Index: Treatment non-response (no improvement-reliable change index) Follow-up: mean 26 weeks	Study population 500 per 1000 Moderate 500 per 1000	575 per 1000 (380 to 865) 575 per 1000 (380 to 865)	RR 1.15 (0.76 to 1.73)	80 (1 study)	⊕⊖⊖⊖ very low1,2,3	
General mental health treatment non-response Post-treatment – available case analysis SCL-90: Global Severity Index: Treatment non-response (no improvement-reliable change index) Follow-up: mean 26 weeks	Study population 459 per 1000 Moderate 460 per 1000	551 per 1000 (354 to 868) 552 per 1000 (354 to 869)	RR 1.2 (0.77 to 1.89)	75 (1 study)	⊕⊖⊖⊖ very low1,2,3	
General mental health mean scores (lower better) Post-treatment – available case analysis SCL-90: Global Severity Index		The mean general mental health mean scores (lower better) post-treatment – available case analysis in the intervention groups was 0.24 standard		75 (1 study)	⊕⊖⊖⊖ very low1,3,4	SMD -0.24 (-0.7 to 0.21)

Follow-up: mean 26 weeks	deviations lower (0.7 lower to 0.21 higher)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Risk of bias due to statistically significant group differences at baseline
2 Total number of events is less than 300 (a threshold rule-of-thumb)
3 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
4 Total population size is less than 400 (a threshold rule-of-thumb)

General mental health outcomes: co-parenting intervention versus enhanced treatment as usual

There was single study (N=28) evidence for a moderate benefit of a co-parenting intervention on reducing psychological distress (p=0.09). However, confidence in this effect estimate is low due to very serious imprecision as a result of the very small sample size and the 95% CI includes both no effect and appreciable benefit (Table 186).

Table 186: Summary of findings table for effects of co-parenting intervention compared with enhanced treatment as usual on general mental health outcomes

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	General mental health: Co-parenting intervention versus Enhanced TAU				
Psychological distress mean scores post-treatment – available case analysis Keller Symptom Questionnaire: Psychological		The mean psychological distress mean scores post-treatment – available case analysis in the intervention groups was 0.65 standard deviations lower		28 (1 study)	⊕⊕⊖⊖ low1,2	SMD -0.65 (-1.42 to 0.11)

distress Follow-up: mean 6 weeks	(1.42 lower to 0.11 higher)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total population size is less than 400 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

7.5.12 Clinical evidence for effects on mother–infant attachment (by intervention)

Summary of findings can be found in the tables presented in this section. The full GRADE evidence profiles and associated forest plots can be found in Appendix 22 and Appendix 19, respectively.

Mother-infant attachment: structured psychological interventions (CBT or IPT) versus treatment as usual or enhanced treatment as usual

There was high to very low quality evidence from up to two studies for moderate to large benefits of structured psychological interventions (CBT or IPT) in reducing mother–infant attachment problems at endpoint (p=0.01-0.003) and at long-term follow-up (p=0.16-0.35), mean mother–infant attachment scores (p=0.20), mother–infant play frequency (p <0.00001), and maternal sensitivity (p=0.10). There was, however, no evidence for clinically or statistically significant benefits on mother–infant behaviour management problems (p=0.53-0.56) or mother–infant attachment mean scores at short-term follow-up (p=0.29), and although there was a statistically significant effect of CBT/IPT on exclusive breastfeeding at 6 months, the effect size was too small to be considered clinically meaningful (p=0.02-0.03) (Table 187).

Table 187: Summary of findings table for effects of mother–infant relationship interventions compared with treatment as usual or enhanced treatment as usual on mother–infant attachment outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Mother-infant attachment: Structured psychological interventions (CBT or IPT) versus TAU/Enhanced TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Mother-infant attachment problems Post-treatment – ITT analysis Maternal report: Mother-infant relationship problems Follow-up: mean 20 weeks	Study population 827 per 1000	537 per 1000 (405 to 719)	RR 0.65 (0.49 to 0.87)	102 (1 study)	⊕⊕⊖⊖ low1	
	Moderate 827 per 1000	538 per 1000 (405 to 719)				
Mother-infant attachment problems Post-treatment – available case analysis Maternal report: Mother-infant relationship problems Follow-up: mean 20 weeks	Study population 743 per 1000	468 per 1000 (319 to 676)	RR 0.63 (0.43 to 0.91)	78 (1 study)	⊕⊕⊖⊖ low1	
	Moderate 743 per 1000	468 per 1000 (319 to 676)				
Mother-infant attachment mean score Post-treatment – available case analysis Prenatal Attachment Inventory or Maternal Attachment Inventory		The mean mother–infant attachment mean score post-treatment – available case analysis in the intervention groups was 2.28 standard deviations higher (1.17 lower to 5.73 higher)		76 (2 studies)	⊕⊖⊖⊖ very low2,3,4	SMD 2.28 (-1.17 to 5.73)

Follow-up: 8-15 weeks						
Mother-infant play frequency Post-treatment – ITT analysis Mother-infant interaction: Play frequency (Events were mother played with infant once or more every day) Follow-up: mean 52 weeks	Study population 339 per 1000	535 per 1000 (457 to 623)	RR 1.58 (1.35 to 1.84)	903 (1 study)	⊕⊕⊕⊕ high	
	Moderate 339 per 1000	536 per 1000 (458 to 624)				
Mother-infant play frequency Post-treatment – available case analysis Mother-infant interaction: Play frequency (Events were mother played with infant once or more every day) Follow-up: mean 52 weeks	Study population 432 per 1000	687 per 1000 (596 to 790)	RR 1.59 (1.38 to 1.83)	705 (1 study)	⊕⊕⊕⊕ high	
	Moderate 432 per 1000	687 per 1000 (596 to 791)				
Maternal sensitivity mean scores post-treatment – available case analysis Study-specific task: Attention bias for distressed infant faces reaction time paradigm Follow-up: mean 15 weeks		The mean maternal sensitivity mean scores post-treatment – available case analysis in the intervention groups was 0.86 standard deviations higher (0.16 lower to 1.88 higher)		17 (1 study)	⊕⊖⊖⊖ very low3,4,5,6	SMD 0.86 (-0.16 to 1.88)
Mother-infant behaviour management problems Post-treatment – ITT analysis Maternal report: Behaviour management problems Follow-up: mean 20 weeks	Study population 577 per 1000	519 per 1000 (363 to 738)	RR 0.9 (0.63 to 1.28)	102 (1 study)	⊕⊕⊖⊖ low1,4	
	Moderate 577 per 1000	519 per 1000 (364 to 739)				
Mother-infant behaviour management	Study population 371 per 1000	442 per 1000 (256 to 761)	RR 1.19 (0.69 to 2.05)	78 (1 study)	⊕⊕⊖⊖ low1,4	

problems Post-treatment – available case analysis Maternal report: Behaviour management problems Follow-up: mean 20 weeks	Moderate 371 per 1000	441 per 1000 (256 to 761)				
Discontinued (exclusive) breastfeeding <6 months – ITT analysis Infant feeding-no longer exclusively breastfeeding by 26 weeks Follow-up: mean 52 weeks	Study population 909 per 1000 Moderate 909 per 1000	864 per 1000 (827 to 909) 864 per 1000 (827 to 909)	RR 0.95 (0.91 to 1)	903 (1 study)	⊕⊕⊕⊕ high	
Discontinued (exclusive) breastfeeding <6 months Post-treatment – available case analysis Infant feeding-no longer exclusively breastfeeding by 26 weeks Follow-up: mean 52 weeks	Study population 889 per 1000 Moderate 889 per 1000	826 per 1000 (782 to 880) 827 per 1000 (782 to 880)	RR 0.93 (0.88 to 0.99)	727 (1 study)	⊕⊕⊕⊕ high	
Mother-infant attachment mean scores Short follow-up (9-16 weeks post-intervention) – available case analysis Maternal Attachment Inventory Follow-up: mean 21 weeks		The mean mother–infant attachment mean scores short follow-up (9-16 weeks post-intervention) – available case analysis in the intervention groups was 0.32 standard deviations higher (0.27 lower to 0.91 higher)		45 (1 study)	⊕⊕⊖⊖ low3,4	SMD 0.32 (-0.27 to 0.91)
Mother-infant attachment problems Long follow-up (25-103 weeks post-intervention) – ITT analysis Maternal report:	Study population 481 per 1000 Moderate 481 per 1000	620 per 1000 (433 to 885) 620 per 1000 (433 to 885)	RR 1.29 (0.9 to 1.84)	102 (1 study)	⊕⊕⊖⊖ low1,4	

Mother-infant relationship problems Follow-up: mean 78 weeks					
Mother-infant attachment problems Long follow-up (25-103 weeks post-intervention) – available case analysis Maternal report: Mother-infant relationship problems Follow-up: mean 78 weeks	Study population 426 per 1000	523 per 1000 (336 to 817)	RR 1.23 (0.79 to 1.92)	87 (1 study)	⊕⊕⊖⊖ low1,4
	Moderate 426 per 1000	524 per 1000 (337 to 818)			

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total number of events is less than 300 (a threshold rule-of-thumb)
2 There is evidence of considerable heterogeneity of study effect sizes
3 Total population size is less than 400 (a threshold rule-of-thumb)
4 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
5 Risk of bias due to unclear blinding of outcome assessment
6 Paper omits data

Mother-infant attachment: facilitated self-help versus treatment as usual

There was no evidence for a clinically or statistically significant benefit (p=0.12) of facilitated self-help on maternal attitude towards motherhood (Table 188).

Table 188: Summary of findings table for effects of facilitated self-help compared with treatment as usual on mother–infant attachment outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Mother-infant attachment: Facilitated self-help versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Maternal attitude towards motherhood mean scores post-treatment – available case analysis Postnatal Bonding Questionnaire Follow-up: mean 17 weeks		The mean maternal attitude towards motherhood mean scores post-treatment – available case analysis in the intervention groups was 0.41 standard deviations higher (0.11 lower to 0.92 higher)		59 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	SMD 0.41 (-0.11 to 0.92)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total population size is less than 400 (a threshold rule-of-thumb)
[2] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
[3] Paper omits data

Mother-infant attachment: listening visits versus treatment as usual

There was low quality, single study evidence for moderate benefits of listening visits on reducing mother–infant attachment problems (p=0.01-0.06) and behaviour management problems (p=0.12 for ITT analysis). However, the effect on behaviour management problems was not clinically or statistically significant when using an available case analysis approach (p=0.84) and effects on mother–infant attachment problems were not maintained at long-term follow-up (p=0.69-0.89). There were also no clinically or statistically significant effects of listening visits on breastfeeding discontinuation before 6 months (p=0.33-0.36) (Table 189).

Table 189: Summary of findings table for effects of listening visits compared with treatment as usual on mother–infant attachment outcomes

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Mother-infant attachment: Listening visits versus TAU				
Mother-infant attachment problems Post-treatment – ITT analysis Maternal report: Mother-infant relationship problems Follow-up: mean 20 weeks	Study population		RR 0.71 (0.54 to 0.92)	100 (1 study)	⊕⊕⊖⊖ low1	
	827 per 1000	587 per 1000 (447 to 761)				
	Moderate					
	827 per 1000	587 per 1000 (447 to 761)				
Mother-infant attachment problems Post-treatment – available case analysis Maternal report: Mother-infant relationship problems Follow-up: mean 20 weeks	Study population		RR 0.72 (0.51 to 1.01)	78 (1 study)	⊕⊕⊖⊖ low1,2	
	743 per 1000	535 per 1000 (379 to 750)				
	Moderate					
	743 per 1000	535 per 1000 (379 to 750)				
Mother-infant behaviour management problems Post-treatment – ITT analysis Maternal report: Behaviour management problems Follow-up: mean 20 weeks	Study population		RR 0.72 (0.48 to 1.09)	100 (1 study)	⊕⊕⊖⊖ low1,2	
	577 per 1000	415 per 1000 (277 to 629)				
	Moderate					
	577 per 1000	415 per 1000 (277 to 629)				
Mother-infant behaviour management problems Post-treatment – available case analysis Maternal report: Behaviour management problems Follow-up: mean 20 weeks	Study population		RR 0.94 (0.52 to 1.7)	78 (1 study)	⊕⊕⊖⊖ low1,2	
	371 per 1000	349 per 1000 (193 to 631)				
	Moderate					
	371 per 1000	349 per 1000 (193 to 631)				
	Study population					

Discontinued breastfeeding <6 months – ITT analysis Infant feeding-breast feeding stopped by 26 weeks Follow-up: mean 52 weeks	383 per 1000	422 per 1000 (345 to 514)	RR 1.1 (0.9 to 1.34)	731 (1 study)	⊕⊕⊖⊖ low1,2
	Moderate				
	383 per 1000	421 per 1000 (345 to 513)			
Discontinued breastfeeding <6 months Post-treatment – available case analysis Infant feeding-breast feeding stopped by 26 weeks Follow-up: mean 52 weeks	Study population		RR 1.09 (0.91 to 1.3)	557 (1 study)	⊕⊕⊖⊖ low1,2
	504 per 1000	549 per 1000 (458 to 655)			
	Moderate				
	504 per 1000	549 per 1000 (459 to 655)			
Mother-infant attachment problems Long follow-up (25-103 weeks post-intervention) – ITT analysis Maternal report: Mother-infant relationship problems Follow-up: mean 78 weeks	Study population		RR 1.08 (0.73 to 1.6)	100 (1 study)	⊕⊕⊖⊖ low1,2
	481 per 1000	519 per 1000 (351 to 769)			
	Moderate				
	481 per 1000	519 per 1000 (351 to 770)			
Mother-infant attachment problems Long follow-up (25-103 weeks post-intervention) – available case analysis Maternal report: Mother-infant relationship problems Follow-up: mean 78 weeks	Study population		RR 0.96 (0.58 to 1.59)	86 (1 study)	⊕⊕⊖⊖ low1,2
	426 per 1000	409 per 1000 (247 to 677)			
	Moderate				
	426 per 1000	409 per 1000 (247 to 677)			

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total number of events is less than 300 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Mother-infant attachment: social support versus treatment as usual

There were no clinically or statistically significant (p=0.13-0.55) benefits of social support for positive mother–infant feeding or teaching interactions (Table 190).

Table 190: Summary of findings table for effects of social support compared with treatment as usual on mother–infant attachment outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Mother-infant attachment: Social support versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Mother-infant feeding interaction Post-treatment – available case analysis Nursing Child Assessment Satellite Training Scale (NCAST): Feeding Follow-up: mean 12 weeks		The mean mother–infant feeding interaction post-treatment – available case analysis in the intervention groups was 0.18 standard deviations lower (0.79 lower to 0.42 higher)		43 (1 study)	⊕⊕⊖⊖ low1,2	SMD -0.18 (-0.79 to 0.42)
Mother-infant teaching interaction Post-treatment – available case analysis NCAST: Teaching Follow-up: mean 12 weeks		The mean mother–infant teaching interaction post-treatment – available case analysis in the intervention groups was 0.45 standard deviations lower (1.04 lower to 0.13 higher)		46 (1 study)	⊕⊕⊖⊖ low1,2	SMD -0.45 (-1.04 to 0.13)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total population size is less than 400 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Mother-infant attachment: psychologically (CBT/IPT)-informed psychoeducation versus enhanced treatment as usual

There was low quality single study (N=194) evidence for a moderate benefit of psychoeducation on maternal sense of competence at post-treatment (p <0.0001), and a small (but not appreciable) benefit maintained at short-term follow-up (p=0.02; Table 191).

Table 191: Summary of findings table for effects of psychologically (CBT/IPT)-informed psychoeducation compared with enhanced treatment as usual on mother–infant attachment outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Mother-infant attachment: Psychologically (CBT/IPT)-informed psychoeducaiton versus Enhanced TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Maternal competence/confidence mean scores post-treatment – available case analysis Parenting Sense of Competence Scale (PSCS): Efficacy Follow-up: mean 6 weeks		The mean maternal competence/confidence mean scores post-treatment – available case analysis in the intervention groups was 0.57 standard deviations higher (0.29 to 0.86 higher)		194 (1 study)	⊕⊕⊖⊖ low1	SMD 0.57 (0.29 to 0.86)
Maternal competence/confidence mean scores Short follow-up (9-16 weeks post-intervention) – available case analysis PSCS: Efficacy Follow-up: mean 13 weeks		The mean maternal competence/confidence mean scores short follow-up (9-16 weeks post-intervention) – available case analysis in the intervention groups was 0.35 standard deviations higher (0.06 to 0.63 higher)		194 (1 study)	⊕⊕⊖⊖ low1	SMD 0.35 (0.06 to 0.63)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total population size is less than 400 (a threshold rule-of-thumb)

Mother-infant attachment: home visits versus treatment as usual

There was no evidence for statistically or clinically significant effects (p=0.23-0.37) of home visits on mother–infant attachment problems (Table 192).

Table 192: Summary of findings table for effects of home visits compared with treatment as usual on mother–infant attachment outcomes

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Mother-infant attachment: Home visits versus TAU/Enhanced TAU				
Mother-infant attachment problems Post-treatment – ITT analysis NCAST ≤35 Follow-up: mean 104 weeks	Study population		RR 0.87 (0.69 to 1.09)	364 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	
	476 per 1000	414 per 1000 (328 to 518)				
	Moderate					
	476 per 1000	414 per 1000 (328 to 519)				
Mother-infant attachment problems Post-treatment – available case analysis NCAST ≤35 Follow-up: mean 104 weeks	Study population		RR 0.79 (0.47 to 1.32)	249 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	
	211 per 1000	167 per 1000 (99 to 279)				
	Moderate					
	211 per 1000	167 per 1000 (99 to 279)				

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the

estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Risk of bias due to statistically significant group differences at baseline
[2] Total number of events is less than 300 (a threshold rule-of-thumb)
[3] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Mother-infant attachment: mother-infant relationship interventions versus treatment as usual or enhanced treatment as usual

There was mixed, but largely non-significant, evidence for the effects of mother–infant relationship interventions on mother–infant attachment outcomes (Table 193). There was very low quality evidence from two studies (N=175) for a moderate benefit of mother–infant relationship interventions on reducing attachment problems (p ≤0.0001). There was also single study (N=75-95) evidence for moderate benefits of mother–infant relationship interventions on maternal sensitivity and maternal structuring treatment response (reliable change index; p=0.46-0.53) and behaviour management problems (for ITT [p=0.04] but not available case [p=0.62] analysis). However, confidence in the effect estimates for the dichotomous measures of maternal sensitivity and structuring were very low due to risk of bias concerns (statistically significant differences in infant age at baseline and selective reporting bias) and very serious imprecision (as the optimal information size of 300 events was not met and the 95% CIs include appreciable harm, no effect and appreciable benefit). There was also low quality single study (N=58-71) evidence for moderate to large benefits of mother–infant relationship interventions on maternal sensitivity (p=0.001), maternal structuring (p=0.02), child responsiveness (p=0.006), and child involvement (p=0.002) at long follow-up (25-103 weeks post-intervention), but not on maternal nonintrusiveness (p=0.15) or maternal nonhostility (p=0.94) at long-term follow-up, or child attachment security at very long-term (>104 weeks post-intervention) follow-up (p=0.11). In addition, evidence from up to four studies (N=146-378) found no evidence for statistically or clinically significant effects on continuous measures of mother–infant attachment or positive interactions (p=0.47), maternal sensitivity (p=0.15), maternal structuring (p=0.13), or child involvement/positive engagement (p=0.22). There was also no evidence for clinically or statistically significant effects on maternal nonintrusiveness (p=0.72-0.76), child responsiveness (p=0.67-0.69) or child involvement (p=0.96-1.00) dichotomous treatment responses, or continuous measures of maternal intrusive behaviour (p=0.16), maternal nonhostility (p=0.67), maternal sense of competence (p=0.55), child responsiveness (p=0.16), or child attachment security (p=0.06) at endpoint, or mother–infant positive interaction, maternal sensitivity or maternal intrusive behaviour mean scores at intermediate follow-up (p=0.46-1.00), or mother–infant attachment problems at long-term follow-up (p=0.30-0.45). Moreover, there was single study evidence for a large harm (p <0.00001) of mother–infant relationship interventions on mother–infant positive interaction mean scores at very

long follow-up with effects favouring enhanced treatment as usual (telephone support).

Table 193: Summary of findings table for effects of mother–infant relationship interventions compared with treatment as usual or enhanced treatment-as-usual on mother–infant attachment outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Mother-infant attachment: Mother-infant relationship interventions versus TAU/Enhanced TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Mother-infant attachment problems Post-treatment – ITT analysis Maternal report: Mother-infant relationship problems or Parent-Infant Relationship Global Assessment Scale (PIR-GAS): Treatment non-response (no improvement-reliable change index) Follow-up: 20-26 weeks	Study population 793 per 1000 Moderate 789 per 1000	436 per 1000 (333 to 571) 434 per 1000 (331 to 568)	RR 0.55 (0.42 to 0.72)	175 (2 studies)	⊕⊖⊖⊖ very low1,2	
Mother-infant attachment problems Post-treatment – available case analysis Maternal report: Mother-infant relationship problems or PIR-GAS: Treatment non-response (no improvement-reliable change index) Follow-up: 20-26 weeks	Study population 736 per 1000 Moderate 736 per 1000	405 per 1000 (302 to 545) 405 per 1000 (302 to 545)	RR 0.55 (0.41 to 0.74)	151 (2 studies)	⊕⊖⊖⊖ very low1,2	
Mother-infant positive interaction mean scores post-treatment – available case analysis Dyadic Mutuality Code, PIR-GAS, Behavioural		The mean mother–infant positive interaction mean scores post-treatment – available case analysis in the intervention groups was		378 (4 studies)	⊕⊖⊖⊖ very low3,4,5	SMD 0.15 (-0.26 to 0.56)

observation: Positive mother–infant interaction or Global Rating Scales of Mother-Infant Interaction: Overall mother–infant interaction Follow-up: 5-26 weeks		0.15 standard deviations higher (0.26 lower to 0.56 higher)			
Maternal sensitivity treatment response Post-treatment – ITT analysis Emotional Availability Scales (EAS): Maternal sensitivity: Treatment response (improvement-reliable change index) Follow-up: mean 26 weeks	Study population 75 per 1000	125 per 1000 (32 to 488)	RR 1.67 (0.43 to 6.51)	80 (1 study)	⊕⊖⊖⊖ very low2,5,6,7
	Moderate 75 per 1000	125 per 1000 (32 to 488)			
Maternal sensitivity treatment response Post-treatment – available case analysis EAS: Maternal sensitivity: Treatment response (improvement-reliable change index) Follow-up: mean 26 weeks	Study population 81 per 1000	131 per 1000 (34 to 512)	RR 1.62 (0.42 to 6.31)	75 (1 study)	⊕⊖⊖⊖ very low2,5,6,7
	Moderate 81 per 1000	131 per 1000 (34 to 511)			
Maternal sensitivity mean scores post-treatment – available case analysis EAS: Maternal sensitivity or Behavioural observation: Maternal sensitivity or Global Rating Scales of Mother-Infant Interaction: Maternal sensitive behaviour Follow-up: 5-28 weeks		The mean maternal sensitivity mean scores post-treatment – available case analysis in the intervention groups was 0.23 standard deviations higher (0.08 lower to 0.53 higher)		332 (4 studies)	⊕⊖⊖⊖ very low4,5,8	SMD 0.23 (-0.08 to 0.53)
Maternal structuring treatment response Post-treatment – ITT analysis EAS: Maternal structuring: Treatment response (improvement-reliable	Study population 100 per 1000	150 per 1000 (46 to 491)	RR 1.5 (0.46 to 4.91)	80 (1 study)	⊕⊖⊖⊖ very low2,5,6,7	
	Moderate 100 per 1000	150 per 1000 (46 to 491)				

change index) Follow-up: mean 26 weeks					
Maternal structuring treatment response Post-treatment – available case analysis EAS: Maternal structuring: Treatment response (improvement-reliable change index) Follow-up: mean 26 weeks	Study population 108 per 1000	158 per 1000 (49 to 515)	RR 1.46 (0.45 to 4.76)	75 (1 study)	⊕⊖⊖⊖ very low2,5,6,7	
	Moderate 108 per 1000	158 per 1000 (49 to 514)				
Maternal structuring mean scores post-treatment – available case analysis EAS: Maternal structuring Follow-up: 26-28 weeks		The mean maternal structuring mean scores post-treatment – available case analysis in the intervention groups was 0.25 standard deviations higher (0.07 lower to 0.58 higher)		146 (2 studies)	⊕⊖⊖⊖ very low4,5,6,7	SMD 0.25 (-0.07 to 0.58)
Maternal nonintrusiveness treatment response Post-treatment – ITT analysis EAS: Maternal nonintrusiveness: Treatment response (improvement-reliable change index) Follow-up: mean 26 weeks	Study population 175 per 1000	151 per 1000 (56 to 408)	RR 0.86 (0.32 to 2.33)	80 (1 study)	⊕⊖⊖⊖ very low2,5,6,7	
	Moderate 175 per 1000	151 per 1000 (56 to 408)				
Maternal nonintrusiveness treatment response Post-treatment – available case analysis EAS: Maternal nonintrusiveness: Treatment response (improvement-reliable change index) Follow-up: mean 26 weeks	Study population 189 per 1000	157 per 1000 (59 to 426)	RR 0.83 (0.31 to 2.25)	75 (1 study)	⊕⊖⊖⊖ very low2,5,6,7	
	Moderate 189 per 1000	157 per 1000 (59 to 425)				
Maternal nonintrusive behaviour mean scores post-treatment – available case analysis EAS: Maternal		The mean maternal nonintrusive behaviour mean scores post-treatment – available case analysis in the intervention groups		146 (2 studies)	⊕⊖⊖⊖ very low4,5,6,7	SMD 0.24 (-0.08 to 0.57)

nonintrusiveness Follow-up: 26-28 weeks		was 0.24 standard deviations higher (0.08 lower to 0.57 higher)			
Maternal intrusive behaviour mean scores post-treatment – available case analysis Global Rating Scales of Mother-Infant Interaction: Maternal intrusive behaviour Follow-up: mean 7 weeks		The mean maternal intrusive behaviour mean scores post-treatment – available case analysis in the intervention groups was 0.28 standard deviations higher (0.11 lower to 0.68 higher)		98 (1 study)	⊕⊕⊖⊖ low4,5	SMD 0.28 (-0.11 to 0.68)
Maternal nonhostility mean scores post-treatment – available case analysis EAS: Maternal nonhostility Follow-up: mean 28 weeks		The mean maternal nonhostility mean scores post-treatment – available case analysis in the intervention groups was 0.1 standard deviations higher (0.37 lower to 0.57 higher)		71 (1 study)	⊕⊖⊖⊖ very low4,5,9	SMD 0.1 (-0.37 to 0.57)
Child responsiveness treatment response Post-treatment – ITT analysis EAS: Child responsiveness: Treatment response (improvement-reliable change index) Follow-up: mean 26 weeks	Study population 100 per 1000 Moderate 100 per 1000	75 per 1000 (18 to 314) 75 per 1000 (18 to 314)	RR 0.75 (0.18 to 3.14)	80 (1 study)	⊕⊖⊖⊖ very low2,5,6,7	
Child responsiveness treatment response Post-treatment – available case analysis EAS: Child responsiveness: Treatment response (improvement-reliable change index) Follow-up: mean 26 weeks	Study population 108 per 1000 Moderate 108 per 1000	79 per 1000 (19 to 329) 79 per 1000 (19 to 328)	RR 0.73 (0.18 to 3.04)	75 (1 study)	⊕⊖⊖⊖ very low2,5,6,7	
Child responsiveness mean scores post-treatment – available case analysis EAS: Child		The mean child responsiveness mean scores post-treatment – available case analysis in the intervention groups was		146 (2 studies)	⊕⊖⊖⊖ very low3,4,5,6,7	SMD 0.38 (-0.15 to 0.92)

responsiveness Follow-up: 26-28 weeks		0.38 standard deviations higher (0.15 lower to 0.92 higher)			
Child involvement treatment response Post-treatment – ITT analysis EAS: Child involvement: Treatment response (improvement-reliable change index) Follow-up: mean 26 weeks	Study population 175 per 1000	175 per 1000 (68 to 453)	RR 1 (0.39 to 2.59)	80 (1 study)	⊕⊖⊖⊖ very low2,5,6,7
	Moderate 175 per 1000	175 per 1000 (68 to 453)			
Child involvement treatment response Post-treatment – available case analysis EAS: Child involvement: Treatment response (improvement-reliable change index) Follow-up: mean 26 weeks	Study population 189 per 1000	184 per 1000 (72 to 473)	RR 0.97 (0.38 to 2.5)	75 (1 study)	⊕⊖⊖⊖ very low2,5,6,7
	Moderate 189 per 1000	183 per 1000 (72 to 472)			
Child involvement/positive engagement mean scores post-treatment – available case analysis EAS: Child involvement or Behavioural observation: Child involvement or Global Rating Scales of Mother-Infant Interaction: Infant positive engagement Follow-up: 5-28 weeks		The mean child involvement/positive engagement mean scores post-treatment – available case analysis in the intervention groups was 0.14 standard deviations higher (0.09 lower to 0.37 higher)		332 (4 studies)	⊕⊕⊕⊖ moderate4	SMD 0.14 (-0.09 to 0.37)
Child attachment security mean scores post-treatment – available case analysis Attachment Q Set (AQS III): Child attachment security Follow-up: mean 57 weeks		The mean child attachment security mean scores post-treatment – available case analysis in the intervention groups was 0.45 standard deviations higher (0.02 lower to 0.93 higher)		71 (1 study)	⊕⊕⊖⊖ low4,5	SMD 0.45 (-0.02 to 0.93)
	Study population				

Mother-infant behaviour management problems Post-treatment – ITT analysis Maternal report: Behaviour management problems Follow-up: mean 20 weeks	577 per 1000 Moderate 577 per 1000	346 per 1000 (219 to 560) 346 per 1000 (219 to 560)	RR 0.6 (0.38 to 0.97)	95 (1 study)	⊕⊕⊖⊖ low2	
Mother-infant behaviour management problems Post-treatment – available case analysis Maternal report: Behaviour management problems Follow-up: mean 20 weeks	Study population 371 per 1000 Moderate 371 per 1000	RR 0.85 (0.46 to 1.59) 316 per 1000 (171 to 591) 315 per 1000 (171 to 590)		76 (1 study)	⊕⊕⊖⊖ low2,5	
Maternal confidence/competence mean scores post-treatment – available case analysis Maternal report: Beliefs about competence Follow-up: mean 25 weeks		The mean maternal confidence/competence mean scores post-treatment – available case analysis in the intervention groups was 0.12 standard deviations lower (0.52 lower to 0.28 higher)		96 (1 study)	⊕⊕⊖⊖ low4,5	SMD -0.12 (-0.52 to 0.28)
Mother-infant positive interaction mean scores Intermediate follow-up (17-24 weeks post-intervention) – available case analysis Global Rating Scales of Mother-Infant Interaction: Overall mother–infant interaction Follow-up: mean 25 weeks		The mean mother–infant positive interaction mean scores intermediate follow-up (17-24 weeks post-intervention) – available case analysis in the intervention groups was 0 standard deviations higher (0.4 lower to 0.4 higher)		96 (1 study)	⊕⊕⊖⊖ low4	SMD 0 (-0.4 to 0.4)
Maternal sensitivity mean scores Intermediate follow-up (17-24 weeks post-intervention) – available case analysis Global Rating Scales of Mother-Infant Interaction: Maternal sensitive behaviour		The mean maternal sensitivity mean scores intermediate follow-up (17-24 weeks post-intervention) – available case analysis in the intervention groups was 0.15 standard deviations higher		96 (1 study)	⊕⊕⊖⊖ low4,5	SMD 0.15 (-0.25 to 0.55)

Follow-up: mean 25 weeks		(0.25 lower to 0.55 higher)			
Maternal intrusive behaviour mean scores Intermediate follow-up (17-24 weeks post-intervention) – available case analysis Global Rating Scales of Mother-Infant Interaction: Maternal intrusive behaviour Follow-up: mean 25 weeks		The mean maternal intrusive behaviour mean scores intermediate follow-up (17-24 weeks post-intervention) – available case analysis in the intervention groups was 0.13 standard deviations higher (0.27 lower to 0.53 higher)		96 (1 study)	⊕⊕⊖⊖ low4,5	SMD 0.13 (-0.27 to 0.53)
Mother-infant attachment problems Long follow-up (25-103 weeks post-intervention) – ITT analysis Maternal report: Mother-infant relationship problems Follow-up: mean 78 weeks	Study population 481 per 1000	558 per 1000 (380 to 822)	RR 1.16 (0.79 to 1.71)	95 (1 study)	⊕⊕⊖⊖ low2,5	
	Moderate 481 per 1000	558 per 1000 (380 to 823)				
Mother-infant attachment problems Long follow-up (25-103 weeks post-intervention) – available case Maternal report: Mother-infant relationship problems Follow-up: mean 78 weeks	Study population 426 per 1000	536 per 1000 (345 to 830)	RR 1.26 (0.81 to 1.95)	88 (1 study)	⊕⊕⊖⊖ low2,5	
	Moderate 426 per 1000	537 per 1000 (345 to 831)				
Maternal sensitivity mean scores Long follow-up (25-103 weeks post-intervention)- Available case analysis EAS: Maternal sensitivity Follow-up: mean 57 weeks		The mean maternal sensitivity mean scores long follow-up (25-103 weeks post-intervention)- available case analysis in the intervention groups was 0.81 standard deviations higher (0.33 to 1.3 higher)		71 (1 study)	⊕⊕⊖⊖ low4	SMD 0.81 (0.33 to 1.3)
Maternal structuring mean scores Long follow-up (25-103 weeks post-intervention) – available case analysis		The mean maternal structuring mean scores long follow-up (25-103 weeks post-intervention) – available case analysis		71 (1 study)	⊕⊕⊖⊖ low4	SMD 0.56 (0.09 to 1.03)

EAS: Maternal structuring Follow-up: mean 57 weeks	in the intervention groups was 0.56 standard deviations higher (0.09 to 1.03 higher)			
Maternal nonintrusive behaviour mean scores Long follow-up (25-103 weeks post-intervention) – available case analysis EAS: Maternal nonintrusiveness Follow-up: mean 57 weeks	The mean maternal nonintrusive behaviour mean scores long follow-up (25-103 weeks post-intervention) – available case analysis in the intervention groups was 0.34 standard deviations higher (0.13 lower to 0.81 higher)	71 (1 study)	⊕⊕⊖⊖ low4,5	SMD 0.34 (-0.13 to 0.81)
Maternal nonhostility mean scores Long follow-up (25-103 weeks post-intervention) – available case analysis EAS: Maternal nonhostility Follow-up: mean 57 weeks	The mean maternal nonhostility mean scores long follow-up (25-103 weeks post-intervention) – available case analysis in the intervention groups was 0.02 standard deviations lower (0.48 lower to 0.45 higher)	71 (1 study)	⊕⊕⊖⊖ low4	SMD -0.02 (-0.48 to 0.45)
Child responsiveness mean scores Long follow-up (25-103 weeks post-intervention) – available case analysis EAS: Child responsiveness Follow-up: mean 57 weeks	The mean child responsiveness mean scores long follow-up (25-103 weeks post-intervention) – available case analysis in the intervention groups was 0.68 standard deviations higher (0.2 to 1.16 higher)	71 (1 study)	⊕⊕⊖⊖ low4	SMD 0.68 (0.2 to 1.16)
Child involvement mean scores Long follow-up (25-103 weeks post-intervention) – available case analysis EAS: Child involvement Follow-up: mean 57 weeks	The mean child involvement mean scores long follow-up (25-103 weeks post-intervention) – available case analysis in the intervention groups was 0.74 standard deviations higher (0.26 to 1.23 higher)	71 (1 study)	⊕⊕⊖⊖ low4	SMD 0.74 (0.26 to 1.23)
Mother-infant positive interaction mean scores Very long follow-up	The mean mother–infant positive interaction mean scores	58 (1 study)	⊕⊕⊖⊖ low4	SMD -1.82 (-2.44 to -1.2)

(>104 weeks post-intervention) – available case analysis Behavioural observation: Positive mother–infant interaction Follow-up: mean 271 weeks	very long follow-up (>104 weeks post-intervention) – available case analysis in the intervention groups was 1.82 standard deviations lower (2.44 to 1.2 lower)			
Child attachment security mean scores Very long follow-up (>104 weeks post-intervention) – available case analysis Attachment Story Completion Task Follow-up: mean 271 weeks	The mean child attachment security mean scores very long follow-up (>104 weeks post-intervention) – available case analysis in the intervention groups was 0.42 standard deviations higher (0.1 lower to 0.95 higher)	58 (1 study)	⊕⊕⊖⊖ low4,5	SMD 0.42 (-0.1 to 0.95)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Risk of bias due to statistically significant group differences at baseline and non-blind outcome assessment
2 Total number of events is less than 300 (a threshold rule-of-thumb)
3 There is evidence of substantial heterogeneity of study effect sizes
4 Total population size is less than 400 (a threshold rule-of-thumb)
5 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
6 Risk of bias due to statistically significant group differences at baseline
7 Paper omits data
8 There is evidence of moderate heterogeneity of study effect sizes
9 Evidence of selective reporting for this outcome measure

Mother-infant attachment: mother-infant relationship intervention with video feedback versus mother–infant relationship intervention with verbal feedback

A single study compared two mother–infant relationship intervention arms and found no differences in effects on maternal sense of competence or on maternal perceptions of infant behaviour between the intervention arm including video feedback and the intervention arm including verbal feedback (p=0.16-0.58; Table 194).

Table 194: Summary of findings table for effects of mother–infant relationship intervention with video feedback compared with mother–infant relationship intervention with verbal feedback on mother–infant attachment outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Mother-infant attachment: Mother-infant relationship intervention with video feedback versus mother–infant relationship intervention with verbal feedback	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Maternal confidence/competence mean scores post-treatment – available case analysis PSCS Follow-up: mean 3 weeks		The mean maternal confidence/competence mean scores post-treatment – available case analysis in the intervention groups was 0.48 standard deviations lower (1.13 lower to 0.18 higher)		37 (1 study)	⊕⊖⊖⊖ very low1,2,3	SMD -0.48 (-1.13 to 0.18)
Maternal perceptions of infant behaviour mean scores post-treatment – available case analysis Neonatal Perception Inventory: Maternal perceptions of infant behaviour Follow-up: mean 3 weeks		The mean maternal perceptions of infant behaviour mean scores post-treatment – available case analysis in the intervention groups was 0.17 standard deviations higher (0.45 lower to 0.8 higher)		40 (1 study)	⊕⊖⊖⊖ very low1,2,3	SMD 0.17 (-0.45 to 0.8)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total population size is less than 400 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
3 Paper omits data

Mother-infant attachment: mother-infant relationship intervention (and facilitated self-help aimed at the eating disorder) versus listening visits (and facilitated self-help aimed at the eating disorder)

There was very low quality single study (N=80) evidence for moderate to large benefits (Table 195) of a mother–infant relationship intervention relative to listening visits for women with eating disorders for reducing mealtime conflict (p=0.01-0.02), maternal inappropriate verbal responses (p=0.06-0.08), and infant autonomy (p=0.01-0.03), but not for maternal intrusions (p=0.38-0.49).

Table 195: Summary of findings table for effects of mother–infant relationship intervention (+ facilitated self-help) compared with listening visits (+ facilitated self-help) on mother–infant attachment outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Mother-infant attachment: Mother-infant relationship intervention (and facilitated self-help) versus listening visits (and facilitated self-help)	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Mealtime conflict Post-treatment – ITT analysis Behavioural observation of mealtime: Significant mealtime conflict (conflict was judged to have occurred if a conflict was at a severe or marked level of clinical concern [rating of 1 or 2] for any 2-minute observational period) Follow-up: mean 35 weeks	Study population 550 per 1000	275 per 1000 (154 to 489)	RR 0.5 (0.28 to 0.89)	80 (1 study)	⊕⊖⊖⊖ very low1,2	
	Moderate 550 per 1000	275 per 1000 (154 to 489)				
	Study population					

Mealtime conflict Post-treatment – available case analysis Behavioural observation of mealtime: Significant mealtime conflict (conflict was judged to have occurred if a conflict was at a severe or marked level of clinical concern [rating of 1 or 2] for any 2-minute observational period) Follow-up: mean 35 weeks	538 per 1000	237 per 1000 (124 to 447)	RR 0.44 (0.23 to 0.83)	77 (1 study)	⊕⊖⊖⊖ very low1,2
	Moderate				
	539 per 1000	237 per 1000 (124 to 447)			
Maternal inappropriate verbal responses Post-treatment – ITT analysis Behavioural observation of mealtime: Maternal inappropriate verbal responses Follow-up: mean 35 weeks	Study population		RR 0.7 (0.48 to 1.04)	80 (1 study)	⊕⊖⊖⊖ very low1,2,3
	675 per 1000	472 per 1000 (324 to 702)			
	Moderate				
	675 per 1000	472 per 1000 (324 to 702)			
Maternal inappropriate verbal responses Post-treatment – available case analysis Behavioural observation of mealtime: Maternal inappropriate verbal responses Follow-up: mean 35 weeks	Study population		RR 0.67 (0.44 to 1.02)	77 (1 study)	⊕⊖⊖⊖ very low1,2,3
	667 per 1000	447 per 1000 (293 to 680)			
	Moderate				
	667 per 1000	447 per 1000 (293 to 680)			
Maternal intrusions Post-treatment – ITT analysis Behavioural observation of mealtime: Maternal intrusions Follow-up: mean 35 weeks	Study population		RR 0.81 (0.45 to 1.46)	80 (1 study)	⊕⊖⊖⊖ very low1,2,3
	400 per 1000	324 per 1000 (180 to 584)			
	Moderate				
	400 per 1000	324 per 1000 (180 to 584)			
Maternal intrusions Post-treatment – available case analysis Behavioural observation of mealtime: Maternal intrusions Follow-up: mean 35 weeks	Study population		RR 0.75 (0.4 to 1.42)	77 (1 study)	⊕⊖⊖⊖ very low1,2,3
	385 per 1000	288 per 1000 (154 to 546)			
	Moderate				
	385 per 1000	289 per 1000 (154 to 547)			
Infant autonomy Post-treatment – ITT analysis Behavioural observation	Study population		RR 1.36 (1.04 to 1.79)	80 (1 study)	⊕⊖⊖⊖ very low1,2
	625 per 1000	850 per 1000 (650 to 1000)			

of mealtime: Infant autonomy Follow-up: mean 35 weeks	Moderate 625 per 1000 — 850 per 1000 (650 to 1000)			
Infant autonomy Post-treatment – available case analysis Behavioural observation of mealtime: Infant autonomy Follow-up: mean 35 weeks	Study population 641 per 1000 — 897 per 1000 (692 to 1000) Moderate 641 per 1000 — 897 per 1000 (692 to 1000)	RR 1.4 (1.08 to 1.81)	77 (1 study)	⊕⊖⊖⊖ very low1,2,3

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total number of events is less than 300 (a threshold rule-of-thumb)
2 Paper omits data
3 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

7.5.13 Clinical evidence for effects on mental health outcomes (sub-analyses)

Depression outcomes by baseline diagnostic status

There was evidence for statistically significant subgroup differences by baseline diagnostic status for depression diagnosis (ITT analysis [p=0.007]; available case analysis [p=0.03]) with clinically and statistically significant benefits observed for psychosocial interventions on depression diagnosis where the participants had a clinical diagnosis of depression at baseline (usually assessed using a structured psychiatric interview [p <0.00001]), clinically but not statistically significant benefits observed for participants who had baseline symptoms of depression (scored above threshold on a depression rating scale) for ITT analysis or clinically and statistically significant benefits but with a less precise estimate of effect for available case analysis (p=0.008), and no evidence for clinically or statistically significant effects of psychosocial interventions on depression diagnosis for participants with subthreshold symptoms at baseline (p=0.86-0.93).

Depression outcomes by format

There was evidence for statistically significant subgroup differences by format for mean depression symptoms (ITT analysis [p=0.03]) with large benefits of psychosocial interventions delivered in an individual format on mean depression

symptoms (p=0.01) but no evidence for clinically or statistically significant benefits of group psychosocial interventions on mean depression symptoms (p=0.65).

Depression outcomes by treatment timing, mode of delivery and intensity

There were no clinically meaningful subgroup differences for the sub-analyses of depression outcomes by treatment timing (for instance, antenatal, postnatal, antenatal and postnatal), mode of delivery (for instance, face-to-face, telephone, internet), or intensity (high [>16 sessions of contact with healthcare professional], moderate [8-16 sessions of contact with healthcare professional]; low [<8 sessions of contact with healthcare professional]).

Sub-analyses for other outcomes

There was insufficient data to enable sub-analysis by baseline diagnosis status, treatment timing, mode of delivery, format or intensity for anxiety, adjustment disorder, PTSD, OCD, general mental health, or mother–infant attachment outcomes.

7.5.14 Clinical evidence for effects of interventions aimed at substance or alcohol misuse

Alcohol use during pregnancy: brief alcohol reduction intervention versus alcohol assessment only

As reviewed in STADE2009B, there was single study evidence (N=142) for a statistically significant effect of a brief alcohol reduction intervention on the number if women who remained abstinent throughout the trial (p=0.04). However, the effect size was small and did not reach the threshold for appreciable clinical benefit (RR 1.20 [1.01, 1.42]). Moreover, there were no clinically or statistically significant treatment effects on the number of women who were abstinent following the trial (RR 1.11 [0.93, 1.33]; p=0.25) or the number of antenatal drinking episodes (SMD -0.20 [-0.45, 0.05]; p=0.12).

Alcohol use during pregnancy: brief cognitive behavioural intervention versus usual advice

As reviewed in STADE2009B, there was single study evidence (N=72) for a moderate effect of a brief cognitive behavioural intervention on the number of women abstaining from alcohol at follow-up (RR 1.25 [0.97, 1.61]). However, this effect was not statistically significant (p=0.09), and there was no evidence for a clinically or statistically significant effect on the average drinks per month (SMD -0.45 [-0.92, 0.02]; p=0.06).

Alcohol use during pregnancy: motivational interviews versus brief written information

As reported in STADE2009B, there was no evidence from a single study (N=34) for clinically or statistically significant effects of motivational interviews on the total standard units of alcohol (SMD -0.05 [-0.73, 0.62]; p=0.88) or days abstinent (SMD 0.32 [-0.36, 1.00]; p=0.36). Two additional studies which met eligibility criteria for

this review (OSTERMAN2012, OSTERMAN2014) provided consistent results with no clinically or statistically significant benefits of motivational interviews observed on drink days per week (not estimable), drink days per month (SMD 0.03 [-0.37, 0.44]; p=0.87), harmful drinking behaviour/dependency symptoms (SMD 0.10 [-0.31, 0.51]; p=0.64), psychological needs (SMD 0.14 [-0.39, 0.67]; p=0.61), or motivation to decrease alcohol use (SMD -0.03 [-0.35, 0.30]; p=0.88).

Alcohol use during pregnancy: brief intervention versus routine care

As reported in STADE2009B, there was single study (N=255) evidence for a small and statistically significant effect of a brief intervention for alcohol use on abstinence in the third trimester (RR 1.08 [1.02, 1.14]; p=0.01), although this effect failed to reach the threshold for a clinically appreciable benefit. As reported in STADE2009B, there was however evidence for a large, and clinically and statistically significant, effect of this brief intervention on alcohol reduction in the third trimester (SMD -3.09 [-3.46, -2.73]; p <0.00001). Moreover, an additional study (N=179) identified by this review (MARAIS2011) also found evidence for clinically and statistically significant effects of a brief intervention on alcohol reduction in the third trimester (RR 1.74 [1.31, 2.32]; p=0.0001).

Alcohol use in the postnatal period: psychologically-informed psychoeducation versus control

A single study (N=235) which met eligibility criteria for this review but not for any of the Cochrane reviews (FLEMING2008) found no evidence for clinically significant benefits, although some of the effects reached statistical significance, of a psychologically-informed psychoeducational intervention (based on CBT and motivational interviewing principles) for women who screened positively for at-risk drinking in the postnatal period on total number of standard drinks (SMD -0.35 [-0.61, -0.09]; p=0.007), number of drinking days (SMD -0.14 [-0.40, 0.11]; p=0.27), or number of heavy drinking (≥4 drinks) days (SMD -0.34 [-0.59, -0.08]; p=0.01).

Alcohol use in the postnatal period: home visits versus control

As reported in TURNBULL2012 there was no evidence from two studies (N=248) for clinically or statistically significant benefits of home visits in the postnatal period on continued alcohol use (RR 1.08 [0.83, 1.41]; p=0.55).

Illicit drug use during pregnancy: any psychosocial intervention versus control

As reported in TERPLAN2007 and updated with two studies identified by this review (WINHUSEN2008, YONKERS2012), there was no evidence (N=239-822) for any clinically or statistically significant benefits of psychosocial interventions on retention in treatment (RR 1.02 [0.95, 1.09]; p=0.63) or retention at one month or more (RR 1.07 [0.87, 1.33]; p=0.52).

Illicit drug use during pregnancy: manual-based interventions versus control

As reported in TERPLAN2007, there was no evidence from three studies (N=226) for a clinically or statistically significant effect of manual-based interventions on retention in treatment (RR 0.93 [0.81, 1.06]; p=0.27).

Illicit drug use during pregnancy: contingency management versus control

As reported in TERPLAN2007, there was no evidence from four studies (N=213) for a clinically or statistically significant effect of contingency management on retention in treatment (RR 1.14 [0.98, 1.34]; p=0.09).

Illicit drug use in the postnatal period: contingency management versus control

A long-term follow-up (SILVERMAN2002) of a study included in TERPLAN2007 (Silverman et al., 2001) met the eligibility criteria for this review but not for any of the Cochrane reviews and provided single study (N=40) evidence for a large benefit of contingency management on continued illicit drug abstinence at three year follow-up (RR 5.00 [0.64, 39.06]; p=0.12). However, this effect estimate was imprecise (with a very small sample size and the 95% CI including both no effect and a measure of appreciable benefit) and not statistically significant.

Illicit drug use in the postnatal period: home visits versus control

As reported in TURNBULL2012 there was no evidence from two studies (N=248) for clinically or statistically significant benefits of home visits in the postnatal period on continued illicit drug use (RR 0.95 [0.75, 1.20]; p=0.64). There was evidence from two studies ([N=211] reported in TURNBULL2012) for a large effect of postnatal home visits (in favour of the intervention) on failure to enrol in a drug treatment programme, however, this effect was not statistically significant and there was considerable heterogeneity between effect estimates (RR 0.45 [0.10, 1.94]; p=0.28). There was single study (N=103) evidence (reported in TURNBULL2012) for a moderate, and clinically and statistically significant, benefit of postnatal home visits on failure to remain in drug treatment at 4 weeks (RR 0.54 [0.35, 0.84]; p=0.007). However, this effect was not maintained at 90 days (RR 0.93 [0.69, 1.25]; p=0.63).

Illicit drug use in the postnatal period: self-help versus attention-placebo control

A single study (N=143) which met eligibility criteria for this review but not for any of the Cochrane reviews (ONDERSMA2014) found evidence for a large, and clinically and statistically significant benefit, of self-help on illicit drug abstinence at 13-week follow-up (RR 2.68 [1.20, 5.97]; p=0.02). Moreover, a moderate and clinically significant benefit was maintained at 26-week follow-up (RR 1.41 [0.57, 3.49]; p=0.46), although this effect estimate was imprecise and failed to reach statistical significance.

Depression in the postnatal period: psychologically-informed psychoeducation versus control

A single study (N=205) which met eligibility criteria for this review but not for any of the Cochrane reviews (FLEMING2008) found no evidence for a clinically or statistically significant benefit of a psychologically-informed psychoeducational intervention for women who screened positive for at-risk drinking in the postnatal period on depression at 6-month follow-up (SMD -0.22 [-0.50, 0.05]; p=0.11).

Mother-infant attachment: home visits versus control

As reported in TURNBULL2012 there was no evidence from a single study (N=124) for a clinically or statistically significant benefit of postnatal home visits on the number of women who discontinued breastfeeding before six months (RR 1.00 [0.81, 1.23]; p=1.00).

7.5.15 Clinical evidence for effects on quality of life (by intervention)

Summary of findings can be found in the tables presented in this section. The full GRADE evidence profiles and associated forest plots can be found in Appendix 22 and Appendix 19, respectively.

Quality of life: structured psychological interventions (CBT or IPT) versus treatment as usual or enhanced treatment as usual

There was high quality evidence from three studies (N=897) for a moderate benefit of CBT or IPT on social support at post-treatment when an available case analysis was used (p <0.00001). However, the effect estimate from the ITT analysis of a single study (N=93) failed to meet clinical or statistical significance thresholds (p=0.07), though this could be a consequence of a lack of power. Conversely at short-term follow-up, there was single study (N=93) low quality evidence for a moderate benefit of CBT (and home visits) relative to home visits-only on social support using an ITT analysis approach (p=0.003), however, the available case analysis of another single study (N=45) found no evidence for clinically or statistically significant effects of IPT relative to treatment as usual on social support at short-term follow-up (p=0.34) (Table 196).

There was single study (N=212) low quality evidence for a moderate benefit of CBT relative to treatment as usual on maternal stress (p=0.0001). However, the confidence in this effect estimate was downgraded as the rule-of-thumb threshold for optimal information size (that is, 400 participants) was not met and there was a high risk of selective reporting bias. The same study (N=284) also found evidence for a small effect of CBT relative to treatment as usual on wellbeing (p=0.0005), however, this effect estimate did not meet the criteria for a clinically meaningful and appreciable benefit (as SMD<0.5) (Table 196).

There was single study (N=284) low quality evidence for a small benefit of CBT relative to treatment as usual on functional impairment (p=0.0009), however, again

despite statistical significance, the threshold for clinical significance was not reached. Very low quality evidence from four studies (although only two studies included in each analysis [N=146-897]) found no evidence for clinically or statistically significant effects of CBT or IPT relative to treatment as usual or enhanced treatment as usual on life functioning at post-treatment using an available case analysis approach (p=0.91) or an ITT analysis (p=0.70). However, there was single study (N=93) low quality evidence for a moderate benefit of CBT (and home visits) relative to home visits-only on life functioning at short-term follow-up using an ITT analysis approach (p=0.005) (Table 196).

Table 196: Summary of findings table for effects of structured psychological interventions (CBT or IPT) compared with treatment as usual or enhanced treatment as usual on quality of life outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Quality of life: Structured psychological interventions (CBT or IPT) versus TAU/Enhanced TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Social support Post-treatment (mean score at endpoint or first measurement) – ITT analysis Interpersonal Support Evaluation List (ISEL) Follow-up: mean 15 weeks		The mean social support post-treatment (mean score at endpoint or first measurement) – ITT analysis in the intervention groups was 0.38 standard deviations higher (0.03 lower to 0.79 higher)		93 (1 study)	⊕⊕⊖⊖ low1,2	SMD 0.38 (-0.03 to 0.79)
Social support Post-treatment (mean score at endpoint or first measurement or change score) – available case analysis Social Provision Scale (SPS): Social support or ISEL or Multidimensional Scale for Perceived Social Support		The mean social support post-treatment (mean score at endpoint or first measurement or change score) – available case analysis in the intervention groups was 0.63 standard deviations higher (0.5 to 0.77 higher)		897 (3 studies)	⊕⊕⊕⊕ high	SMD 0.63 (0.5 to 0.77)

Follow-up: 12-52 weeks				
Life functioning Post-treatment (mean score at endpoint or first measurement) – ITT analysis Global Assessment of Functioning Scale or Social Adjustment Scale (SAS): Social and leisure domain Follow-up: 15-44 weeks	The mean life functioning post-treatment (mean score at endpoint or first measurement) – ITT analysis in the intervention groups was 0.44 standard deviations lower (2.65 lower to 1.78 higher)	146 (2 studies)	⊕⊖⊖⊖ very low1,2,3	SMD -0.44 (-2.65 to 1.78)
Life functioning Post-treatment (mean score at endpoint or first measurement) – available case analysis SAS or Global Assessment of Functioning Scale Follow-up: 12-52 weeks	The mean life functioning post-treatment (mean score at endpoint or first measurement) – available case analysis in the intervention groups was 0.1 standard deviations lower (1.92 lower to 1.72 higher)	897 (2 studies)	⊕⊖⊖⊖ very low2,3	SMD -0.1 (-1.92 to 1.72)
Functional impairment Post-treatment (mean score at endpoint or first measurement) – available case analysis CORE-OM: Life functioning Follow-up: mean 26 weeks	The mean functional impairment post-treatment (mean score at endpoint or first measurement) – available case analysis in the intervention groups was 0.4 standard deviations lower (0.63 to 0.16 lower)	284 (1 study)	⊕⊕⊖⊖ low1,4	SMD -0.4 (-0.63 to -0.16)
Parental stress Post-treatment (mean score at endpoint or first measurement or change score) – available case analysis PSI Follow-up: mean 26 weeks	The mean parental stress post-treatment (mean score at endpoint or first measurement or change score) – available case analysis in the intervention groups was 0.53 standard deviations higher (0.26 to 0.81 higher)	212 (1 study)	⊕⊕⊖⊖ low1,4	SMD 0.53 (0.26 to 0.81)
Wellbeing Post-treatment (mean score at endpoint or first measurement) – available case analysis CORE-OM: Well-	The mean wellbeing post-treatment (mean score at endpoint or first measurement) – available case analysis in the intervention	284 (1 study)	⊕⊕⊖⊖ low1,4	SMD -0.42 (-0.65 to -0.18)

being Follow-up: mean 26 weeks	groups was 0.42 standard deviations lower (0.65 to 0.18 lower)			
Social support Short follow-up (mean score at 9-16 week follow-up) – ITT analysis ISEL Follow-up: mean 28 weeks	The mean social support short follow-up (mean score at 9-16 week follow-up) – ITT analysis in the intervention groups was 0.64 standard deviations higher (0.22 to 1.06 higher)	93 (1 study)	⊕⊕⊖⊖ low1	SMD 0.64 (0.22 to 1.06)
Social support Short follow-up (mean score at 9-16 week follow-up) – available case analysis ISEL Follow-up: mean 21 weeks	The mean social support short follow-up (mean score at 9-16 week follow-up) – available case analysis in the intervention groups was 0.29 standard deviations higher (0.3 lower to 0.88 higher)	45 (1 study)	⊕⊕⊖⊖ low1,2	SMD 0.29 (-0.3 to 0.88)
Life functioning Short follow-up (mean score at 9-16 week follow-up) – ITT analysis Global Assessment of Functioning Scale Follow-up: mean 28 weeks	The mean life functioning short follow-up (mean score at 9-16 week follow-up) – ITT analysis in the intervention groups was 0.6 standard deviations higher (0.18 to 1.02 higher)	93 (1 study)	⊕⊕⊖⊖ low1	SMD 0.6 (0.18 to 1.02)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total population size is less than 400 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
3 There was evidence of considerable heterogeneity between effect sizes
4 Paper omits data

Quality of life: IPT versus support group

A single study (N=44) found no evidence for a clinically or statistically significant benefit of IPT relative to a support group on maternal stress as measured by comparing cortisol levels (p=0.14) (Table 197).

Table 197: Summary of findings table for effects of IPT compared with support group on quality of life outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Quality of life: IPT versus support group	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Maternal stress Post-treatment (mean score at endpoint or first measurement) – available case analysis Maternal cortisol levels Follow-up: mean 12 weeks		The mean maternal stress post-treatment (mean score at endpoint or first measurement) – available case analysis in the intervention groups was 0.45 standard deviations lower (1.05 lower to 0.15 higher)		44 (1 study)	⊕⊖⊖⊖ very low1,2,3	SMD -0.45 (-1.05 to 0.15)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 High risk of selection bias due to unclear allocation concealment and statistically significant baseline differences with the control group showing a higher SES score/lower income and higher depression (CES-D) mean score
2 Total population size is less than 400 (a threshold rule-of-thumb)
3 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Quality of life: facilitated self-help versus treatment as usual

There was single study (N=59-143) very low quality evidence for moderate to large benefits of facilitated self-help relative to treatment as usual on social support (p=0.05), functional impairment (p=0.03), and maternal stress using either an ITT (p=0.02) or available case (p=0.02) analysis approach (Table 198).

Table 198: Summary of findings table for effects of facilitated self-help compared with treatment as usual on quality of life outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Quality of life: Facilitated self-help versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Social support Post-treatment (mean score at endpoint or first measurement or change score) – available case analysis SPS: Social support Follow-up: mean 17 weeks		The mean social support post-treatment (mean score at endpoint or first measurement or change score) – available case analysis in the intervention groups was 0.51 standard deviations higher (0.01 lower to 1.03 higher)		59 (1 study)	⊕⊖⊖⊖ very low1,2,3	SMD 0.51 (-0.01 to 1.03)
Functional impairment Post-treatment (mean score at endpoint or first measurement) – available analysis Work and Social Adjustment Scale: Functional impairment Follow-up: mean 17 weeks		The mean functional impairment post-treatment (mean score at endpoint or first measurement) – available analysis in the intervention groups was 0.57 standard deviations lower (1.1 to 0.05 lower)		59 (1 study)	⊕⊖⊖⊖ very low1,3	SMD -0.57 (-1.1 to -0.05)
Parental stress Post-treatment (symptomatology at endpoint or first measurement) – ITT analysis PSI=>260 Follow-up: mean 20 weeks	Study population 611 per 1000 Moderate 611 per 1000	409 per 1000 (293 to 568) 409 per 1000 (293 to 568)	RR 0.67 (0.48 to 0.93)	143 (1 study)	⊕⊖⊖⊖ very low1,3,4	
Parental stress Post-treatment (symptomatology at endpoint or first	Study population 282 per 1000 Moderate	68 per 1000 (20 to 223)	RR 0.24 (0.07 to 0.79)	84 (1 study)	⊕⊖⊖⊖ very low3,4	

measurement) – available case analysis PSI=>260 Follow-up: mean 20 weeks	282 per 1000	68 per 1000 (20 to 223)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total population size is less than 400 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
3 Paper omits data
4 Total number of events is less than 300 (a threshold rule-of-thumb)

Quality of life: listening visits versus treatment as usual

There was single study (N=277) low quality evidence for small and statistically significant benefits of listening visits on functional impairment (p=0.002) and wellbeing mean scores (p=0.0006), although these effect estimates do not meet criteria for clinical significance (as SMD<0.5). There was also very low quality evidence from another single study (N=41) for a moderate benefit of listening visits on the number of women reporting improvements in wellbeing (p=0.06). However, conversely there was low quality single study (N=211) evidence for a small but statistically significant harm associated with listening visits with higher mean maternal stress scores observed in the intervention group relative to women who received treatment as usual (p=0.001) (Table 199).

Table 199: Summary of findings table for effects of listening visits compared with treatment as usual on quality of life outcomes

Outcomes	**Illustrative comparative risks* (95% CI)** **Assumed risk** **Control**	**Corresponding risk** **Quality of life: Listening visits versus TAU**	**Relative effect (95% CI)**	**No. of Participants (studies)**	**Quality of the evidence (GRADE)**	**Comments**
Functional impairment Post-treatment (mean		The mean functional impairment post-treatment (mean score		277 (1 study)	⊕⊕⊖⊖ low1,2	SMD -0.37 (-0.61 to -0.14)

score at endpoint or first measurement) – available case analysis CORE-OM: Life functioning Follow-up: mean 26 weeks		at endpoint or first measurement) – available case analysis in the intervention groups was 0.37 standard deviations lower (0.61 to 0.14 lower)			
Parental stress Post-treatment (mean score at endpoint or first measurement or change score) – available case analysis PSI Follow-up: mean 26 weeks		The mean parental stress post-treatment (mean score at endpoint or first measurement or change score) – available case analysis in the intervention groups was 0.45 standard deviations higher (0.18 to 0.72 higher)	211 (1 study)	⊕⊕⊖⊖ low1,2	SMD 0.45 (0.18 to 0.72)
Wellbeing Post-treatment (improved wellbeing at endpoint or first measurement) – available case analysis Maternal report: Improvements in wellbeing Follow-up: mean 7 weeks	Study population 571 per 1000 Moderate 571 per 1000	851 per 1000 (560 to 1000) 851 per 1000 (560 to 1000)	RR 1.49 (0.98 to 2.25)	41 (1 study)	⊕⊖⊖⊖ very low2,3,4	
Wellbeing Post-treatment (mean score at endpoint or first measurement) – available case analysis CORE-OM: Well-being Follow-up: mean 26 weeks		The mean wellbeing post-treatment (mean score at endpoint or first measurement) – available case analysis in the intervention groups was 0.42 standard deviations lower (0.66 to 0.18 lower)	277 (1 study)	⊕⊕⊖⊖ low1,2	SMD -0.42 (-0.66 to -0.18)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total population size is less than 400 (a threshold rule-of-thumb)
2 Paper omits data

3 Total number of events is less than 300 (a threshold rule-of-thumb)
4 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Quality of life: directive counselling versus treatment as usual

There was single study (N=90) low quality evidence for a moderate benefit of directive counselling relative to treatment as usual on social support (p=0.05) (Table 200).

Table 200: Summary of findings table for effects of directive counselling compared with treatment as usual on quality of life outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Quality of life: Directive counselling versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Social support Post-treatment (mean score at endpoint or first measurement or change score) – available case analysis SPS: Social support Follow-up: mean 12 weeks		The mean social support post-treatment (mean score at endpoint or first measurement or change score) – available case analysis in the intervention groups was 0.53 standard deviations higher (0.01 to 1.06 higher)		90 (1 study)	⊕⊕⊖⊖ low1	SMD 0.53 (0.01 to 1.06)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total population size is less than 400 (a threshold rule-of-thumb)

Quality of life: post-miscarriage counselling versus treatment as usual

A single study (N=15-19) found evidence for a moderate benefit of post-miscarriage counselling relative to treatment as usual on functional impairment using an available case analysis approach (p=0.21). However, the effect estimate from the ITT

analysis did not meet criteria for clinical or statistical significance (p=0.42). Moreover, confidence in these effect estimates was very low due to risk of bias concerns (statistically significant group difference at baseline) and very serious imprecision (Table 201).

Table 201: Summary of findings table for effects of post-miscarriage counselling compared with treatment as usual on quality of life outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Quality of life: post-miscarriage counselling versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Functional impairment Post-treatment (mean score at endpoint or first measurement) – ITT analysis Short Form Health Survey - 36 (SF-36): Role functioning (sum of role limitation-emotional and social functioning subscales) Follow-up: mean 7 weeks		The mean functional impairment post-treatment (mean score at endpoint or first measurement) – ITT analysis in the intervention groups was 0.37 standard deviations lower (1.28 lower to 0.54 higher)		19 (1 study)	⊕⊖⊖⊖ very low1,2,3	SMD -0.37 (-1.28 to 0.54)
Functional impairment Post-treatment (mean score at endpoint or first measurement) – available case analysis SF-36: Role functioning (sum of role limitation-emotional and social functioning subscales) Follow-up: mean 7 weeks		The mean functional impairment post-treatment (mean score at endpoint or first measurement) – available case analysis in the intervention groups was 0.68 standard deviations lower (1.73 lower to 0.37 higher)		15 (1 study)	⊕⊖⊖⊖ very low1,2,3	SMD -0.68 (-1.73 to 0.37)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 High risk of selection bias due to unclear allocation concealment and statistically significant baseline differences between groups in ethnicity (80% Hispanic in intervention group and 44% in TAU) and Hispanic ethnicity was associated with primary outcome with higher depression scores in Hispanic group
2 Total population size is less than 400 (a threshold rule-of-thumb)
3 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Quality of life: post-traumatic birth counselling versus treatment as usual

There was single study (N=103) low quality evidence for a large benefit of post-traumatic birth counselling relative to treatment as usual on maternal stress symptomatology (p=0.04) (Table 202).

Table 202: Summary of findings table for effects of post-traumatic birth counselling compared with treatment as usual on quality of life outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Quality of life: post-traumatic birth counselling versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Parental stress Post-treatment (symptomatology at endpoint or first measurement) – ITT analysis Depression Anxiety Stress Scale (DASS): Stress>19 Follow-up: mean 13 weeks	Study population 321 per 1000 Moderate 321 per 1000	141 per 1000 (64 to 308) 141 per 1000 (64 to 308)	RR 0.44 (0.2 to 0.96)	103 (1 study)	⊕⊕⊖⊖ low1	
Parental stress Post-treatment (symptomatology at endpoint or first	Study population 321 per 1000 Moderate	141 per 1000 (64 to 308)	RR 0.44 (0.2 to 0.96)	103 (1 study)	⊕⊕⊖⊖ low1	

measurement) – available case analysis Depression Anxiety Stress Scale (DASS): Stress>19 Follow-up: mean 13 weeks	321 per 1000	141 per 1000 (64 to 308)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total number of events is less than 300 (a threshold rule-of-thumb)

Quality of life: social support versus treatment as usual

High to very low quality evidence from up to two studies (N=30-653) found no evidence for clinically or statistically significant effects of social support relative to treatment as usual on social support (p=0.93), maternal cortisol levels (p=0.53), self-esteem (p=0.48), or loneliness at post-treatment (p=0.29) or short-term follow-up (p=0.18). There was low quality evidence from two studies (N=101) for a small and statistically significant benefit of social support on maternal stress (p=0.03), however, this effect estimate did not meet criteria for a clinically meaningful and appreciable benefit (as SMD<0.5) (Table 203).

Table 203: Summary of findings table for effects of social support compared with treatment as usual on quality of life outcomes

Outcomes	**Illustrative comparative risks* (95% CI)**		**Relative effect (95% CI)**	**No. of Participants (studies)**	**Quality of the evidence (GRADE)**	**Comments**
	Assumed risk	**Corresponding risk**				
	Control	**Quality of life: Social support versus TAU**				
Social support Post-treatment (mean score at endpoint or first measurement or change score) – available case analysis ISEL or SPS: Social support		The mean social support post-treatment (mean score at endpoint or first measurement or change score) – available case analysis in the intervention groups was 0.04 standard		111 (2 studies)	⊕⊖⊖⊖ very low1,2,3	SMD 0.04 (-0.87 to 0.96)

Follow-up: 12-14 weeks	deviations higher (0.87 lower to 0.96 higher)			
Parental stress Post-treatment (mean score at endpoint or first measurement or change score) – available case analysis Perceived Stress Scale or Child-Care Stress Checklist Follow-up: 8-14 weeks	The mean parental stress post-treatment (mean score at endpoint or first measurement or change score) – available case analysis in the intervention groups was 0.43 standard deviations lower (0.83 to 0.04 lower)	101 (2 studies)	⊕⊕⊖⊖ low2	SMD -0.43 (-0.83 to -0.04)
Maternal cortisol levels Post-treatment (mean score at endpoint or first measurement) – available case analysis Follow-up: mean 12 weeks	The mean maternal cortisol levels post-treatment (mean score at endpoint or first measurement) – available case analysis in the intervention groups was 0.23 standard deviations higher (0.49 lower to 0.95 higher)	30 (1 study)	⊕⊕⊖⊖ low2,3	SMD 0.23 (-0.49 to 0.95)
Self-esteem Post-treatment (mean score at endpoint or first measurement or change score) – available case analysis Coopersmith's Self-Esteem Inventory or SES Follow-up: 8-14 weeks	The mean self-esteem post-treatment (mean score at endpoint or first measurement or change score) – available case analysis in the intervention groups was 0.14 standard deviations higher (0.25 lower to 0.53 higher)	101 (2 studies)	⊕⊕⊖⊖ low2,3	SMD 0.14 (-0.25 to 0.53)
Loneliness Post-treatment (mean score at endpoint or first measurement) – available case analysis UCLA Loneliness Scale Follow-up: 8-12 weeks	The mean loneliness post-treatment (mean score at endpoint or first measurement) – available case analysis in the intervention groups was 0.26 standard deviations lower (0.74 lower to 0.22 higher)	653 (2 studies)	⊕⊕⊖⊖ low3,4	SMD -0.26 (-0.74 to 0.22)
Loneliness Short follow-up (mean score at 9-16 week follow-up) – available case	The mean loneliness short follow-up (mean score at 9-16 week follow-up) – available case analysis in the	600 (1 study)	⊕⊕⊕⊕ high	SMD -0.11 (-0.27 to 0.05)

analysis UCLA Loneliness Scale Follow-up: mean 24 weeks	intervention groups was 0.11 standard deviations lower (0.27 lower to 0.05 higher)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 There was evidence of considerable heterogeneity between effect sizes
2 Total population size is less than 400 (a threshold rule-of-thumb)
3 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
4 There was evidence of moderate heterogeneity between effect sizes

Quality of life: psychologically (CBT/IPT)-informed psychoeducation versus treatment as usual or enhanced treatment as usual

There was single study (N=194) low quality evidence for a moderate benefit of IPT-informed psychoeducation relative to enhanced treatment as usual (non-mental health-focused education and support group) on social support (p <0.00001) at post-treatment, and a small and statistically significant (although no longer clinically meaningful) benefit was maintained at short-term follow-up (p=0.02) (Table 204).

There was also very low quality evidence from two studies (N=128) for a small and statistically significant benefit of CBT- or IPT- informed psychoeducation relative to treatment as usual on functional impairment (p=0.01) at post-treatment (Table 204). However, this effect estimate did not meet criteria for clinical significance (as SMD <0.5). In addition, a single study (N=42) found no evidence for clinically or statistically significant effects of CBT-informed psychoeducation relative to treatment as usual on functional impairment at short-term follow-up (p=0.17).

No evidence was found for clinically or statistically significant effects of psychologically-informed psychoeducation on maternal stress assessed through self-report scales at post-treatment (using an ITT analysis [K=1; N=156; p=0.26] or available case analysis [K=2; N=95; p=0.83]), intermediate follow-up (using an ITT analysis [K=1; N=156; p=0.59] or available case analysis [K=1; N=42; p=0.60]) or long-term follow-up (using an available case analysis [K=1; N=46; p=0.68]). There was also no evidence from a single study (N=53) for clinically or statistically significant effects of CBT-informed psychoeducation relative to treatment as usual on maternal cortisol levels at post-treatment (K=1; N=53; p=0.18). This study (N=46) did find evidence for a moderate benefit at long-term follow-up (p=0.08). However,

confidence in this effect estimate was very low due to statistically significant group differences in this outcome measure at baseline (high risk of selection bias), a high risk of selective reporting bias, and very serious imprecision (Table 204).

A single study (N=156) found no evidence for clinically or statistically significant effects of IPT-informed psychoeducation relative to treatment as usual on happiness at post-treatment (p=0.76) or long-term follow-up (p=0.26) (Table 204).

Table 204: Summary of findings table for effects of psychologically (CBT/IPT)-informed psychoeducation compared with treatment as usual or enhanced treatment as usual on quality of life outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Quality of life: Psychologically (CBT/IPT)-informed psychoeducation versus TAU/Enhanced TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Social support Post-treatment (mean score at endpoint or first measurement) – ITT analysis Perceived Social Support Scale Follow-up: mean 6 weeks		The mean social support post-treatment (mean score at endpoint or first measurement) – ITT analysis in the intervention groups was 0.74 standard deviations higher (0.45 to 1.03 higher)		194 (1 study)	⊕⊕⊖⊖ low1	SMD 0.74 (0.45 to 1.03)
Functional impairment Post-treatment (mean score at endpoint or first measurement) – available case analysis SAS or LIFE: Range of Impaired Functioning Tool Follow-up: mean 13 weeks		The mean functional impairment post-treatment (mean score at endpoint or first measurement) – available case analysis in the intervention groups was 0.46 standard deviations lower (0.81 to 0.1 lower)		128 (2 studies)	⊕⊖⊖⊖ very low1,2	SMD -0.46 (-0.81 to -0.1)
Parental stress Post-treatment (mean score at endpoint or first measurement) – ITT		The mean parental stress post-treatment (mean score at endpoint or first		156 (1 study)	⊕⊕⊖⊖ low1,3	SMD -0.18 (-0.5 to 0.13)

analysis Perceived Stress Scale Follow-up: mean 4 weeks	measurement) – ITT analysis in the intervention groups was 0.18 standard deviations lower (0.5 lower to 0.13 higher)			
Parental stress Post-treatment (mean score at endpoint or first measurement or change score) – available case analysis VAS: Maternal stress or Perceived Stress Scale Follow-up: 13-49 weeks	The mean parental stress post-treatment (mean score at endpoint or first measurement or change score) – available case analysis in the intervention groups was 0.13 standard deviations lower (1.33 lower to 1.07 higher)	95 (2 studies)	⊕⊖⊖⊖ very low1,3,4,5,6	SMD -0.13 (-1.33 to 1.07)
Maternal cortisol levels Post-treatment (mean score at endpoint or first measurement) – available case analysis Average (morning/evening) cortisol (log scores) Follow-up: mean 49 weeks	The mean maternal cortisol levels post-treatment (mean score at endpoint or first measurement) – available case analysis in the intervention groups was 0.37 standard deviations higher (0.17 lower to 0.92 higher)	53 (1 study)	⊕⊖⊖⊖ very low1,3,4,6	SMD 0.37 (-0.17 to 0.92)
Happiness Post-treatment (mean score at endpoint or first measurement) – ITT analysis Subjective Happiness Scale Follow-up: mean 4 weeks	The mean happiness post-treatment (mean score at endpoint or first measurement) – ITT analysis in the intervention groups was 0.05 standard deviations higher (0.27 lower to 0.36 higher)	156 (1 study)	⊕⊕⊖⊖ low1	SMD 0.05 (-0.27 to 0.36)
Social support Short follow-up (mean score at 9-16 week follow-up) – ITT analysis Perceived Social Support Scale Follow-up: mean 13 weeks	The mean social support short follow-up (mean score at 9-16 week follow-up) – ITT analysis in the intervention groups was 0.33 standard	194 (1 study)	⊕⊕⊖⊖ low1	SMD 0.33 (0.05 to 0.62)

	deviations higher (0.05 to 0.62 higher)			
Functional impairment Intermediate follow-up (mean score at 17-24 week follow-up) – available case analysis SAS Follow-up: mean 26 weeks	The mean functional impairment intermediate follow-up (mean score at 17-24 week follow-up) – available case analysis in the intervention groups was 0.43 standard deviations lower (1.05 lower to 0.18 higher)	42 (1 study)	⊕⊕⊖⊖ low1,3	SMD -0.43 (-1.05 to 0.18)
Parental stress Intermediate follow-up (mean score at 17-24 week follow-up) – ITT analysis Perceived Stress Scale Follow-up: mean 26 weeks	The mean parental stress intermediate follow-up (mean score at 17-24 week follow-up) – ITT analysis in the intervention groups was 0.09 standard deviations lower (0.4 lower to 0.23 higher)	156 (1 study)	⊕⊕⊖⊖ low1	SMD -0.09 (-0.4 to 0.23)
Parental stress Intermediate follow-up (mean score at 17-24 week follow-up) – available case analysis Perceived Stress Scale Follow-up: mean 26 weeks	The mean parental stress intermediate follow-up (mean score at 17-24 week follow-up) – available case analysis in the intervention groups was 0.16 standard deviations lower (0.77 lower to 0.45 higher)	42 (1 study)	⊕⊕⊖⊖ low1,3	SMD -0.16 (-0.77 to 0.45)
Happiness Intermediate follow-up (mean score at 17-24 week follow-up) – ITT analysis Subjective Happiness Scale Follow-up: mean 26 weeks	The mean happiness intermediate follow-up (mean score at 17-24 week follow-up) – ITT analysis in the intervention groups was 0.18 standard deviations higher (0.13 lower to 0.5 higher)	156 (1 study)	⊕⊕⊖⊖ low1,3	SMD 0.18 (-0.13 to 0.5)
Parental stress Long follow-up (mean score at >24 week follow-up) – available case analysis	The mean parental stress long follow-up (mean score at >24 week follow-up) – available case	46 (1 study)	⊕⊖⊖⊖ very low1,3,4,6	SMD 0.12 (-0.46 to 0.7)

VAS: Maternal stress Follow-up: mean 101 weeks	analysis in the intervention groups was 0.12 standard deviations higher (0.46 lower to 0.7 higher)			
Maternal cortisol levels Long follow-up (mean score at >24 week follow-up) – available case analysis Average (morning/evening) cortisol (log scores) Follow-up: mean 101 weeks	The mean maternal cortisol levels long follow-up (mean score at >24 week follow-up) – available case analysis in the intervention groups was 0.52 standard deviations lower (1.11 lower to 0.07 higher)	46 (1 study)	⊕⊖⊖⊖ very low1,3,4,6	SMD -0.52 (-1.11 to 0.07)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total population size is less than 400 (a threshold rule-of-thumb)
2 Unclear risk of selection bias as insufficient detail reported with regards to randomisation method and allocation concealment and unclear risk of detection bias as blinding of outcome assessment is not reported
3 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
4 High risk of selection bias due to statistically significant baseline/mid-treatment difference in average maternal salivary cortisol levels (0.62 in intervention group and 0.75 in control group)
5 There was evidence of considerable heterogeneity between effect sizes
6 Papers omit data

Quality of life: home visits versus treatment as usual or enhanced treatment as usual

There was no evidence for clinically or statistically significant effects of home visits relative to treatment as usual or enhanced treatment as usual on a dichotomous measure of maternal stress (using an ITT [K=1; N=364; p=0.34] or available case [K=1; N=249; p=0.59] analysis approach) or on mean maternal stress scores (K=2; N=595; p=0.62) (Table 205).

Table 205: Summary of findings table for effects of home visits compared with treatment as usual or enhanced treatment as usual on quality of life outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Quality of life: Home visits versus TAU/Enhanced TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Parental stress Post-treatment (symptomatology at endpoint or first measurement) – ITT analysis PSI: Severe parenting stress (as defined by Abidin) Follow-up: mean 104 weeks	Study population 389 per 1000	342 per 1000 (261 to 448)	RR 0.88 (0.67 to 1.15)	364 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	
	Moderate 389 per 1000	342 per 1000 (261 to 447)				
Parental stress Post-treatment (symptomatology at endpoint or first measurement) – available case analysis PSI: Severe parenting stress (as defined by Abidin) Follow-up: mean 104 weeks	Study population 81 per 1000	63 per 1000 (26 to 155)	RR 0.78 (0.32 to 1.91)	249 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	
	Moderate 81 per 1000	63 per 1000 (26 to 155)				
Parental stress Post-treatment (mean score at endpoint or first measurement or change score) – available case analysis PSI or Perceived Stress Scale Follow-up: mean 52 weeks		The mean parental stress post-treatment (mean score at endpoint or first measurement or change score) – available case analysis in the intervention groups was 0.06 standard deviations lower (0.29 lower to 0.18 higher)		595 (2 studies)	⊕⊕⊕⊖ moderate[4]	SMD -0.06 (-0.29 to 0.18)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.

Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] High risk of selection bias due to unclear allocation concealment and statistically significant baseline differences in poor psychological resources (37% intervention group versus 50% control) and in prenatal enrolment (41% intervention group and 53% control)
[2] Total number of events is less than 300 (a threshold rule-of-thumb)
[3] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
[4] Paper omits data

Quality of life: mother-infant relationship interventions versus treatment as usual or enhanced treatment as usual

There was no evidence for clinically or statistically significant effects of mother–infant relationship interventions on a dichotomous measure of maternal stress (using an ITT [K=1; N=80; p=0.13] or available case [K=1; N=75; p=0.14] analysis approach) or on mean maternal stress scores (K=2; N=173; p=0.70) (Table 206).

Table 206: Summary of findings table for effects of mother–infant relationship interventions compared with treatment as usual or enhanced treatment as usual on quality of life outcomes

Outcomes	**Illustrative comparative risks* (95% CI)** **Assumed risk** / **Corresponding risk** **Control** / **Quality of life: Mother-infant relationship interventions versus TAU/Enhanced TAU**	**Relative effect (95% CI)**	**No. of Participants (studies)**	**Quality of the evidence (GRADE)**	**Comments**
Parental stress Post-treatment (symptomatology at endpoint or first measurement) – ITT analysis PSI: Treatment non-response (no improvement-reliable change index) Follow-up: mean 26 weeks	Study population 825 per 1000 / 677 per 1000 (520 to 874) Moderate 825 per 1000 / 677 per 1000 (520 to 874)	RR 0.82 (0.63 to 1.06)	80 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	
	Study population				

Parental stress Post-treatment (symptomatology at endpoint or first measurement) – available case analysis PSI: Treatment non-response (no improvement-reliable change index) Follow-up: mean 26 weeks	811 per 1000	657 per 1000 (503 to 868)				
	Moderate					
	811 per 1000	657 per 1000 (503 to 868)	RR 0.81 (0.62 to 1.07)	75 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	
Parental stress Post-treatment (mean score at endpoint or first measurement or change score) – available case analysis PSI or PSS-NICU: Parental role restriction Follow-up: 4-26 weeks		The mean parental stress post-treatment (mean score at endpoint or first measurement or change score) – available case analysis in the intervention groups was 0.06 standard deviations lower (0.36 lower to 0.24 higher)		173 (2 studies)	⊕⊕⊖⊖ low[4]	SMD -0.06 (-0.36 to 0.24)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] High risk of selection bias due to a statistically significant baseline difference in the age of infants (4.4 months old in intervention group versus 5.9 months old in TAU group)
[2] Total number of events is less than 300 (a threshold rule-of-thumb)
[3] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
[4] Total population size is less than 400 (a threshold rule-of-thumb)

Quality of life: psychosomatic intervention versus treatment as usual

A single study (N=127) found no evidence for clinically or statistically significant effects of a psychosomatic intervention relative to treatment as usual on poor social support (p=0.30) or maternal stress (p=0.54) (Table 207).

Table 207: Summary of findings table for effects of a psychosomatic intervention compared with treatment as usual on quality of life outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Quality of life: Psychosomatic intervention versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Poor social support mean scores post-treatment – available case analysis Functional Social Support Questionnaire (FSSQ): Lack of social support Follow-up: mean 34 weeks		The mean poor social support mean scores post-treatment – available case analysis in the intervention groups was 0.18 standard deviations lower (0.53 lower to 0.17 higher)		127 (1 study)	⊕⊖⊖⊖ very low1,2,3	SMD -0.18 (-0.53 to 0.17)
Parental stress mean scores post-treatment – available case analysis Stress Events Scale (Holmes & Rahe, 1967): Stress score value Follow-up: mean 34 weeks		The mean parental stress mean scores post-treatment – available case analysis in the intervention groups was 0.11 standard deviations lower (0.46 lower to 0.24 higher)		127 (1 study)	⊕⊖⊖⊖ very low1,2	SMD -0.11 (-0.46 to 0.24)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Risk of attrition bias due to statistically significant higher drop-out in the control group
2 Total population size is less than 400 (a threshold rule-of-thumb)
3 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Quality of life: mindfulness training versus treatment as usual or enhanced treatment as usual

Single study analyses of data from two studies (N=31/47) found no evidence for clinically or statistically significant effects of mindfulness training relative to waitlist control or enhanced treatment as usual (non-mental health-focused education and support [book]) on maternal stress (p=0.46-0.60) or positive affect (p=0.23) (Table 208).

Table 208: Summary of findings table for effects of mindfulness training compared with treatment as usual or enhanced treatment as usual on quality of life outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Quality of life: Mindfulness training versus Enhanced TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Parental stress Post-treatment (mean score at endpoint or first measurement) – ITT analysis Perceived Stress Scale Follow-up: mean 6 weeks		The mean parental stress post-treatment (mean score at endpoint or first measurement) – ITT analysis in the intervention groups was 0.22 standard deviations higher (0.36 lower to 0.79 higher)		47 (1 study)	⊕⊕⊖⊖ low1,2	SMD 0.22 (-0.36 to 0.79)
Parental stress Post-treatment (mean score at endpoint or first measurement) – available case analysis Perceived Stress Scale Follow-up: mean 10 weeks		The mean parental stress post-treatment (mean score at endpoint or first measurement) – available case analysis in the intervention groups was 0.19 standard deviations lower (0.91 lower to 0.52 higher)		31 (1 study)	⊕⊖⊖⊖ very low1,2,3	SMD -0.19 (-0.91 to 0.52)
Positive affect Post-treatment (mean score at endpoint or first measurement) – available case analysis Positive and Negative Affect		The mean positive affect post-treatment (mean score at endpoint or first measurement) – available case analysis in the intervention groups was 0.44 standard		31 (1 study)	⊕⊖⊖⊖ very low1,2,3	SMD 0.44 (-0.28 to 1.16)

Schedule-Extended: Positive affect Follow-up: mean 10 weeks		deviations higher (0.28 lower to 1.16 higher)	

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total population size is less than 400 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
3 Paper omits data

7.5.16 Clinical evidence for effects on service utilisation (by intervention)

Summary of findings can be found in the tables presented in this section. The full GRADE evidence profiles and associated forest plots can be found in Appendix 22 and Appendix 19, respectively.

Service utilisation: structured psychological interventions (CBT or IPT) versus treatment as usual or enhanced treatment as usual

A single study (N=46-57) found low quality evidence for reduced use of psychotherapy (p=0.06-0.15) and counselling (p=0.05-0.10) associated with IPT relative to treatment as usual and increased use of alternative therapies relative to treatment as usual (p=0.44-0.46). However, confidence in all these effect estimates is low due to very serious imprecision (very small sample size and wide 95% CIs). This study found no evidence for clinically or statistically significant effects of IPT relative to treatment as usual on health visitor use (p=0.90-1.00), antidepressant use (p=0.77-0.86), or use of a self-help support group (p=0.73-0.92) (Table 209).

Table 209: Summary of findings table for effects of structured psychological interventions (CBT or IPT) compared with treatment as usual or enhanced treatment as usual on service utilisation outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Service utilisation: Structured psychological interventions (CBT or IPT) versus TAU/Enhanced TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Use of NHS health visitor Post-Treatment (service utilisation at endpoint or first measurement) – ITT analysis Maternal and Child Health (MACH) nurse advice Follow-up: mean 21 weeks	Study population 536 per 1000 Moderate 536 per 1000	552 per 1000 (343 to 889) 552 per 1000 (343 to 890)	RR 1.03 (0.64 to 1.66)	57 (1 study)	⊕⊕⊖⊖ low1,2	
Use of NHS health visitor Post-Treatment (service utilisation at endpoint or first measurement) – available case analysis MACH nurse advice Follow-up: mean 21 weeks	Study population 435 per 1000 Moderate 435 per 1000	435 per 1000 (226 to 839) 435 per 1000 (226 to 840)	RR 1 (0.52 to 1.93)	46 (1 study)	⊕⊕⊖⊖ low1,2	
Antidepressant medication Post-Treatment (medication use at endpoint or first measurement) – ITT analysis Antidepressant use Follow-up: mean 21 weeks	Study population 643 per 1000 Moderate 643 per 1000	624 per 1000 (418 to 926) 624 per 1000 (418 to 926)	RR 0.97 (0.65 to 1.44)	57 (1 study)	⊕⊕⊖⊖ low1,2	
Antidepressant medication Post-Treatment (medication use at endpoint or first measurement) – available case analysis Antidepressant use	Study population 565 per 1000 Moderate 565 per 1000	520 per 1000 (305 to 887) 520 per 1000 (305 to 887)	RR 0.92 (0.54 to 1.57)	46 (1 study)	⊕⊕⊖⊖ low1,2	

Follow-up: mean 21 weeks				
Psychotherapy Post-Treatment (service utilisation at endpoint or first measurement) – ITT analysis Follow-up: mean 21 weeks	Study population 464 per 1000 — 274 per 1000 (135 to 562) Moderate 464 per 1000 — 274 per 1000 (135 to 561)	RR 0.59 (0.29 to 1.21)	57 (1 study)	⊕⊕⊖⊖ low1,2
Psychotherapy Post-Treatment (service utilisation at endpoint or first measurement) – available case analysis Follow-up: mean 21 weeks	Study population 348 per 1000 — 87 per 1000 (21 to 365) Moderate 348 per 1000 — 87 per 1000 (21 to 365)	RR 0.25 (0.06 to 1.05)	46 (1 study)	⊕⊕⊖⊖ low1,2
Counselling Post-Treatment (service utilisation at endpoint or first measurement) – ITT analysis Follow-up: mean 21 weeks	Study population 607 per 1000 — 376 per 1000 (219 to 662) Moderate 607 per 1000 — 376 per 1000 (219 to 662)	RR 0.62 (0.36 to 1.09)	57 (1 study)	⊕⊕⊖⊖ low1,2
Counselling Post-Treatment (service utilisation at endpoint or first measurement) – available case analysis Follow-up: mean 21 weeks	Study population 522 per 1000 — 219 per 1000 (89 to 517) Moderate 522 per 1000 — 219 per 1000 (89 to 517)	RR 0.42 (0.17 to 0.99)	46 (1 study)	⊕⊕⊖⊖ low1
Self-help support group Post-Treatment (service utilisation at endpoint or first measurement) – ITT analysis Follow-up: mean 21 weeks	Study population 393 per 1000 — 381 per 1000 (196 to 731) Moderate 393 per 1000 — 381 per 1000 (196 to 731)	RR 0.97 (0.5 to 1.86)	57 (1 study)	⊕⊕⊖⊖ low1,2
Self-help support group Post-Treatment (service utilisation at endpoint or first measurement) – available case analysis Follow-up: mean 21 weeks	Study population 261 per 1000 — 217 per 1000 (78 to 613) Moderate 261 per 1000 — 217 per 1000 (78 to 613)	RR 0.83 (0.3 to 2.35)	46 (1 study)	⊕⊕⊖⊖ low1,2
Alternative therapies Post-Treatment (service utilisation at endpoint or first	Study population 286 per 1000 — 380 per 1000 (180 to 803) Moderate	RR 1.33 (0.63 to 2.81)	57 (1 study)	⊕⊕⊖⊖ low1,2

measurement) – ITT analysis Follow-up: mean 21 weeks	286 per 1000	380 per 1000 (180 to 804)			
Alternative therapies Post-Treatment (service utilisation at endpoint or first measurement) – available case analysis Follow-up: mean 21 weeks	Study population		RR 1.67 (0.45 to 6.17)	46 (1 study)	⊕⊕⊖⊖ low1,2
	130 per 1000	218 per 1000 (59 to 805)			
	Moderate				
	130 per 1000	217 per 1000 (58 to 802)			

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total number of events is less than 300 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Service utilisation: facilitated self-help versus treatment as usual

There was single study (N=57-83) evidence that participants who received facilitated self-help showed less use of the childbirth hospital (p=0.29-0.50) or mental health hospital (p=0.28-0.46) than participants who received treatment as usual. However, confidence in these effect estimates is very low due to very serious imprecision and high risk of selective reporting bias. This study found no clinically or statistically significant effects associated with facilitated self-help on a continuous measure of childbirth hospital usage (p=0.36), the ITT analysis for use of maternal general health hospital (p=0.39), the use of mental health outpatient services (dichotomous ITT analysis [p=0.93]; dichotomous available case analysis [p=0.65]; continuous available case analysis [p=0.08]); the use of health community services (dichotomous ITT analysis [p=0.98]; dichotomous available case analysis [p=0.91]; continuous available case analysis [p=0.71]), or the use of antidepressants (dichotomous ITT analysis [p=0.47]; dichotomous available case analysis [p=0.57]; continuous available case analysis [p=0.59]). Effect estimates could not be calculated for the available case analysis of maternal general health hospital (continuous or dichotomous outcome measures) or use of mental health hospital mean scores due to zero cell counts (Table 210).

Table 210: Summary of findings table for effects of facilitated self-help compared with treatment as usual on service utilisation outcomes

Outcomes	Illustrative comparative risks* (95% CI)	Comments

	Assumed risk Control	Corresponding risk Service utilisation: Facilitated self-help versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	
Use of childbirth hospital Post-Treatment (service utilisation at endpoint) – ITT analysis Adult Service Use Schedule (AD-SUS): Childbirth hospital Follow-up: mean 17 weeks	Study population 405 per 1000 Moderate 405 per 1000	 291 per 1000 (162 to 534) 292 per 1000 (162 to 535)	RR 0.72 (0.4 to 1.32)	83 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	
Use of childbirth hospital Post-Treatment (service utilisation at endpoint) – available case analysis AD-SUS: Childbirth hospital Follow-up: mean 17 weeks	Study population 74 per 1000 Moderate 74 per 1000	 33 per 1000 (3 to 347) 33 per 1000 (3 to 347)	RR 0.45 (0.04 to 4.69)	57 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	
Use of childbirth hospital Post-Treatment (service utilisation at endpoint) – available case analysis AD-SUS: Childbirth hospital Follow-up: mean 17 weeks		The mean use of childbirth hospital post-treatment (service utilisation at endpoint) – available case analysis in the intervention groups was 0.24 standard deviations lower (0.77 lower to 0.28 higher)		57 (1 study)	⊕⊖⊖⊖ very low[2,3,4]	SMD -0.24 (-0.77 to 0.28)
Use of maternal general health hospital Post-Treatment (service utilisation at endpoint) – ITT analysis AD-SUS: Maternal general health hospital Follow-up: mean 17 weeks	Study population 357 per 1000 Moderate 357 per 1000	 268 per 1000 (139 to 514) 268 per 1000 (139 to 514)	RR 0.75 (0.39 to 1.44)	83 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	
Use of maternal general health hospital Post-Treatment (service utilisation at endpoint) – available case analysis	See comment	See comment	Not estimable	57 (1 study)	See comment	

AD-SUS: Maternal general health hospital Follow-up: mean 17 weeks					
Use of maternal general health hospital Post-Treatment (service utilisation at endpoint) – available case analysis AD-SUS: Maternal general health hospital Follow-up: mean 17 weeks	See comment	See comment	Not estimable	57 (1 study)	See comment
Use of mental health hospital Post-Treatment (service utilisation at endpoint) – ITT analysis AD-SUS: Mental health hospital Follow-up: mean 17 weeks	Study population 381 per 1000 Moderate 381 per 1000	267 per 1000 (141 to 507) 267 per 1000 (141 to 507)	RR 0.7 (0.37 to 1.33)	83 (1 study)	⊕⊖⊖⊖ very low[1,2,3]
Use of mental health hospital Post-Treatment (service utilisation at endpoint) – available case analysis AD-SUS: Mental health hospital Follow-up: mean 17 weeks	Study population 37 per 1000 Moderate 37 per 1000	11 per 1000 (0 to 263) 11 per 1000 (0 to 262)	RR 0.3 (0.01 to 7.09)	57 (1 study)	⊕⊖⊖⊖ very low[1,2,3]
Use of mental health hospital Post-Treatment (service utilisation at endpoint) – available case analysis AD-SUS: Mental health hospital Follow-up: mean 17 weeks	See comment	See comment	Not estimable	57 (1 study)	See comment
Use of mental health outpatient Post-Treatment (service utilisation at endpoint) – ITT analysis AD-SUS: Mental health out-patient Follow-up: mean 17 weeks	Study population 619 per 1000 Moderate 619 per 1000	607 per 1000 (433 to 860) 607 per 1000 (433 to 860)	RR 0.98 (0.7 to 1.39)	83 (1 study)	⊕⊖⊖⊖ very low[1,2,3]
	Study population				

Use of mental health outpatient Post-Treatment (service utilisation at endpoint) – available case analysis AD-SUS: Mental health out-patient Follow-up: mean 17 weeks	407 per 1000	469 per 1000 (257 to 847)	RR 1.15 (0.63 to 2.08)	57 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	
	Moderate					
	407 per 1000	468 per 1000 (256 to 847)				
Use of mental health outpatient Post-Treatment (service utilisation at endpoint) – available case analysis AD-SUS: Mental health out-patient Follow-up: mean 17 weeks		The mean use of mental health outpatient post-treatment (service utilisation at endpoint) – available case analysis in the intervention groups was 0.47 standard deviations lower (1 lower to 0.06 higher)		57 (1 study)	⊕⊖⊖⊖ very low[2,3,4]	SMD -0.47 (-1 to 0.06)
Use of health community service Post-Treatment (service utilisation at endpoint) – ITT analysis AD-SUS: Health community service Follow-up: mean 17 weeks	Study population		RR 1 (0.91 to 1.1)	83 (1 study)	⊕⊖⊖⊖ very low[1,3]	
	952 per 1000	952 per 1000 (867 to 1000)				
	Moderate					
	952 per 1000	952 per 1000 (866 to 1000)				
Use of health community service Post-Treatment (service utilisation at endpoint) – available case analysis AD-SUS: Health community service Follow-up: mean 17 weeks	Study population		RR 1.01 (0.87 to 1.16)	57 (1 study)	⊕⊖⊖⊖ very low[1,3]	
	926 per 1000	935 per 1000 (806 to 1000)				
	Moderate					
	926 per 1000	935 per 1000 (806 to 1000)				
Use of health community service Post-Treatment (service utilisation at endpoint) – available case analysis AD-SUS: Health community service Follow-up: mean 17 weeks		The mean use of health community service post-treatment (service utilisation at endpoint) – available case analysis in the intervention groups was 0.1 standard deviations higher (0.42 lower to 0.62 higher)		57 (1 study)	⊕⊖⊖⊖ very low[2,3,4]	SMD 0.1 (-0.42 to 0.62)
Antidepressant medication Post-Treatment (medication	Study population		RR 1.09 (0.86 to 1.38)	83 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	
	738 per 1000	805 per 1000 (635 to 1000)				

use at endpoint or first measurement) – ITT analysis AD-SUS: Antidepressant medication Follow-up: mean 17 weeks	Moderate 738 per 1000	804 per 1000 (635 to 1000)			
Antidepressant medication Post-Treatment (medication use at endpoint or first measurement) – available case analysis AD-SUS: Antidepressant medication Follow-up: mean 17 weeks	Study population 633 per 1000 Moderate 633 per 1000	703 per 1000 (488 to 1000) 703 per 1000 (487 to 1000)	RR 1.11 (0.77 to 1.6)	57 (1 study)	⊕⊖⊖⊖ very low[1,2,3]
Antidepressant medication Post-Treatment (medication use at endpoint) – available case analysis AD-SUS: Antidepressant medication Follow-up: mean 17 weeks		The mean antidepressant medication post-treatment (medication use at endpoint) – available case analysis in the intervention groups was 0.14 standard deviations lower (0.66 lower to 0.38 higher)		57 (1 study)	⊕⊖⊖⊖ very low[2,3,4]	SMD -0.14 (-0.66 to 0.38)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total number of events is less than 300 (a threshold rule-of-thumb)
[2] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
[3] Paper omits data
[4] Total population size is less than 400 (a threshold rule-of-thumb)

Service utilisation: listening visits versus treatment as usual

There was single study evidence (N=601-731) for moderate to large effects of listening visits on service utilisation with listening visits associated with greater usage of NHS health visitor services (p=0.01-0.20) and health visitor telephone contact (p=0.0003-0.08) than treatment as usual. However, it is unclear from the

study whether this service utilisation was independent from the intervention and if not, this may be regarded as more of a compliance measure. This same study found evidence for less use of midwife services associated with listening visits relative to treatment as usual when an available case analysis approach was used (p=0.05), however, effects on midwife usage were not clinically or statistically significant when an ITT analysis approach was adopted (p=0.87). There was also no evidence for clinically or statistically significant effects of listening visits on use of maternal general health hospital (p=0.75-0.77) or use of GP (p=0.72-0.74) (Table 211).

Table 211: Summary of findings table for effects of listening visits compared with treatment as usual on service utilisation outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	 Corresponding risk Service utilisation: Listening visits versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Use of maternal general health hospital Post-Treatment (service utilisation at endpoint) – ITT analysis Health Service Use – Use of hospital doctor in last month Follow-up: mean 52 weeks	Study population 219 per 1000 Moderate 219 per 1000	 208 per 1000 (151 to 287) 208 per 1000 (151 to 287)	RR 0.95 (0.69 to 1.31)	731 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	
Use of maternal general health hospital Post-Treatment (service utilisation at endpoint) – available case analysis Health Service Use – Use of hospital doctor in last month Follow-up: mean 52 weeks	Study population 130 per 1000 Moderate 130 per 1000	 121 per 1000 (75 to 194) 121 per 1000 (75 to 194)	RR 0.93 (0.58 to 1.49)	657 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	
Use of NHS health visitor Post-Treatment (service utilisation at endpoint or first measurement) – ITT analysis Health Service Use – Maternal use of NHS health visitor in last	Study population 131 per 1000 Moderate 131 per 1000	 169 per 1000 (116 to 250) 169 per 1000 (115 to 249)	RR 1.29 (0.88 to 1.9)	731 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	

month Follow-up: mean 52 weeks				
Use of NHS health visitor Post-Treatment (service utilisation at endpoint or first measurement) – available case analysis Health Service Use – Maternal use of NHS health visitor in last month Follow-up: mean 52 weeks	Study population 33 per 1000 — 79 per 1000 (39 to 160) Moderate 33 per 1000 — 80 per 1000 (39 to 163)	RR 2.42 (1.19 to 4.93)	657 (1 study)	⊕⊖⊖⊖ very low[1,3]
Health visitor telephone contact Post-Treatment (service utilisation [in last month] at endpoint) – ITT analysis Health Service Use – Health visitor telephone contact in last month Follow-up: mean 52 weeks	Study population 109 per 1000 — 159 per 1000 (105 to 239) Moderate 110 per 1000 — 160 per 1000 (106 to 240)	RR 1.45 (0.96 to 2.18)	731 (1 study)	⊕⊖⊖⊖ very low[1,2,3]
Health visitor telephone contact Post-Treatment (service utilisation [in last month] at endpoint) – available case analysis Health Service Use – Health visitor telephone contact in last month Follow-up: mean 52 weeks	Study population 8 per 1000 — 67 per 1000 (22 to 207) Moderate 8 per 1000 — 66 per 1000 (21 to 203)	RR 8.2 (2.65 to 25.4)	657 (1 study)	⊕⊖⊖⊖ very low[1,3]
Maternal use of midwife Post-Treatment (service utilisation [in last month] at endpoint) – ITT analysis Health Service Use – Maternal use of midwife in last month Follow-up: mean 78 weeks	Study population 246 per 1000 — 241 per 1000 (180 to 323) Moderate 246 per 1000 — 241 per 1000 (180 to 322)	RR 0.98 (0.73 to 1.31)	731 (1 study)	⊕⊖⊖⊖ very low[1,2,3]
Maternal use of midwife Post-Treatment (service utilisation [in last month] at endpoint) – available case analysis Health Service Use – Maternal use of midwife in last month	Study population 94 per 1000 — 41 per 1000 (18 to 95) Moderate 94 per 1000 — 41 per 1000 (18 to 95)	RR 0.44 (0.19 to 1.01)	601 (1 study)	⊕⊖⊖⊖ very low[1,2,3]

Follow-up: mean 78 weeks				
Use of GP Post-Treatment (service utilisation [in last month] at endpoint) – ITT analysis Health Service Use – Use of GP in last month Follow-up: mean 52 weeks	Study population 502 per 1000	487 per 1000 (411 to 577)	RR 0.97 (0.82 to 1.15)	731 (1 study)	⊕⊕⊕⊖ moderate[3]
	Moderate 502 per 1000	487 per 1000 (412 to 577)			
Use of GP Post-Treatment (service utilisation [in last month] at endpoint) – available case analysis Health Service Use – Use of GP in last month Follow-up: mean 52 weeks	Study population 445 per 1000	432 per 1000 (352 to 525)	RR 0.97 (0.79 to 1.18)	657 (1 study)	⊕⊕⊖⊖ low[1,3]
	Moderate 445 per 1000	432 per 1000 (352 to 525)			

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total number of events is less than 300 (a threshold rule-of-thumb)
[2] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
[3] Paper omits data

Service utilisation: social support versus treatment as usual

A single study (N=600-701) found moderate effects of peer-mediated support with the intervention associated with less antidepressant use at post-treatment (p=0.19) and short-term follow-up (p=0.08). However, using an ITT analysis approach effects on antidepressant usage were not clinically or statistically significant (p=0.45-0.54). The same study also found no evidence for clinically or statistically significant effects of peer-mediated support on a continuous measure of health service usage at post-treatment (p=0.35) or short-term follow-up (p=0.82) (Table 212).

Table 212: Summary of findings table for effects of social support compared with treatment as usual on service utilisation outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Service utilisation: Social support versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Health service use Post-Treatment (service utilisation at endpoint) – available case analysis Health service utilisation and cost of care questionnaire: Health service use Follow-up: mean 12 weeks		The mean health service use post-treatment (service utilisation at endpoint) – available case analysis in the intervention groups was 0.08 standard deviations higher (0.08 lower to 0.23 higher)		612 (1 study)	⊕⊕⊕⊕ high	SMD 0.08 (-0.08 to 0.23)
Antidepressant medication Post-Treatment (medication use at endpoint or first measurement) – ITT analysis Health service utilisation and cost of care questionnaire: Current antidepressant use Follow-up: mean 12 weeks	Study population 159 per 1000 Moderate 159 per 1000	180 per 1000 (130 to 251) 180 per 1000 (130 to 251)	RR 1.13 (0.82 to 1.58)	701 (1 study)	⊕⊕⊖⊖ low1,2	
Antidepressant medication Post-Treatment (medication use at endpoint or first measurement) – available case analysis Health service utilisation and cost of care questionnaire: Current antidepressant use Follow-up: mean 12 weeks	Study population 60 per 1000 Moderate 60 per 1000	37 per 1000 (18 to 77) 37 per 1000 (18 to 76)	RR 0.61 (0.3 to 1.27)	612 (1 study)	⊕⊕⊖⊖ low1,2	

Health service use Short follow-up (service utilisation at 9-16 week follow-up) – available case analysis Health service utilisation and cost of care questionnaire: Health service use Follow-up: mean 24 weeks		The mean health service use short follow-up (service utilisation at 9-16 week follow-up) – available case analysis in the intervention groups was 0.02 standard deviations lower (0.18 lower to 0.14 higher)		600 (1 study)	⊕⊕⊕⊕ high	SMD -0.02 (-0.18 to 0.14)
Antidepressant medication Short follow-up (medication use at 9-16 week follow-up) – ITT analysis Health service utilisation and cost of care questionnaire: Current antidepressant use Follow-up: mean 24 weeks	Study population		RR 1.1 (0.82 to 1.46)	701 (1 study)	⊕⊕⊖⊖ low1,2	
	199 per 1000	219 per 1000 (163 to 290)				
	Moderate					
	199 per 1000	219 per 1000 (163 to 291)				
Antidepressant medication Short follow-up (medication use at 9-16 week follow-up) – available case analysis Health service utilisation and cost of care questionnaire: Current antidepressant use Follow-up: mean 24 weeks	Study population		RR 0.59 (0.33 to 1.07)	600 (1 study)	⊕⊕⊖⊖ low1,2	
	93 per 1000	55 per 1000 (31 to 100)				
	Moderate					
	93 per 1000	55 per 1000 (31 to 100)				

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total number of events is less than 300 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

7.5.17 Clinical evidence for effects on experience of care (by intervention)

The review of qualitative evidence for experience of care is in Chapter 6, however, this section includes any experience of care outcomes reported in the psychosocial treatment RCTs. Summary of findings can be found in the tables presented in this section. The full GRADE evidence profiles and associated forest plots can be found in Appendix 22 and Appendix 19, respectively.

Experience of care: mother-infant relationship interventions versus treatment as usual or enhanced treatment as usual

A single study (N=98) found no evidence for clinically or statistically significant effects of a mother–infant relationship intervention relative to enhanced treatment as usual (non-mental health-focused education and support [booklet about infant care]) on satisfaction with the intervention (p=0.21) or satisfaction with the therapeutic alliance in that the mother felt understood (p=1.00) (Table 213).

Table 213: Summary of findings table for effects of mother–infant relationship interventions compared with treatment as usual or enhanced treatment as usual on experience of care outcomes

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Experience of care: Mother-infant relationship interventions versus TAU/Enhanced TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Satisfaction with intervention Post-treatment (mean score at endpoint or first measurement) – available case analysis Maternal report Follow-up: mean 7 weeks		The mean satisfaction with intervention post-treatment (mean score at endpoint or first measurement) – available case analysis in the intervention groups was 0.25 standard deviations higher (0.14 lower to 0.65 higher)		98 (1 study)	⊕⊕⊖⊖ low1,2	SMD 0.25 (-0.14 to 0.65)
Satisfaction with therapeutic alliance (empathetic) Post-treatment (mean		The mean satisfaction with therapeutic alliance (empathetic) post-treatment (mean		98 (1 study)	⊕⊕⊖⊖ low1	SMD 0 (-0.4 to 0.4)

score at endpoint or first measurement) – available case analysis VAS: Therapeutic alliance (mother felt understood) Follow-up: mean 7 weeks	score at endpoint or first measurement) – available case analysis in the intervention groups was 0 standard deviations higher (0.4 lower to 0.4 higher)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total population size is less than 400 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

7.5.18 Clinical evidence for effects on retention in services and treatment acceptability (by intervention)

Summary of findings can be found in the tables presented in this section. The full GRADE evidence profiles and associated forest plots can be found in Appendix 22 and Appendix 19, respectively.

Retention in services and treatment acceptability (using attrition as a proxy measure): structured psychological interventions (CBT or IPT) versus treatment as usual or enhanced treatment as usual

Twelve studies (N=1,983) found no evidence for clinically or statistically significant effects of structured psychological interventions (CBT or IPT) relative to treatment as usual or enhanced treatment as usual on attrition (p=0.41) (Table 214).

Table 214: Summary of findings table for effects of structured psychological interventions (CBT or IPT) compared with treatment as usual or enhanced treatment as usual on retention in services or treatment acceptability (using attrition as a proxy measure)

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Attrition: Structured psychological interventions (CBT or IPT) versus TAU/Enhanced TAU				
Drop-out Incomplete data at endpoint Follow-up: 6-26 weeks	**Study population**		**RR 1.14** (0.83 to 1.55)	1,983 (12 studies)	⊕⊕⊕⊖ **moderate**[1]	
	156 per 1000	**177 per 1000** (129 to 241)				
	Moderate					
	155 per 1000	**177 per 1000** (129 to 240)				

*The basis for the **assumed risk** (for example, the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% CI) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Retention in services and treatment acceptability (using attrition as a proxy measure): CBT versus relational constructivist therapy

A single study (N=60) found no evidence for a clinically or statistically significant difference between CBT and relational constructivist therapy on attrition (p=0.89) (Table 215).

Table 215: Summary of findings table for effects of CBT compared with relational constructivist therapy on retention in services or treatment acceptability (using attrition as a proxy measure)

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Attrition: CBT versus relational constructivist therapy				
Drop-out Incomplete data at endpoint	Study population		RR 0.88 (0.13 to 5.81)	60 (1 study)	⊕⊕⊖⊖ low1,2	
	71 per 1000	63 per 1000 (9 to 415)				
	Moderate					
	71 per 1000	62 per 1000 (9 to 413)				

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total number of events is less than 300 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Retention in services and treatment acceptability (using attrition as a proxy measure): IPT versus support group

A single study (N=48) found no evidence for a clinically or statistically significant difference between IPT and a support group on attrition (p=1.00) (Table 216).

Table 216: Summary of findings table for effects of IPT compared with support group on retention in services or treatment acceptability (using attrition as a proxy measure)

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	 Corresponding risk Attrition: IPT versus support group	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Drop-out Incomplete data at endpoint Follow-up: mean 12 weeks	Study population		RR 1 (0.15 to 6.53)	48 (1 study)	⊕⊖⊖⊖ very low1,2,3	
	83 per 1000	83 per 1000 (13 to 544)				
	Moderate					
	83 per 1000	83 per 1000 (12 to 542)				

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] High risk of selection bias due to unclear allocation concealment and statistically significant baseline differences with the control group showing a higher SES score/lower income and higher depression (CES-D) mean score
[2] Total number of events is less than 300 (a threshold rule-of-thumb)
[3] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Retention in services and treatment acceptability (using attrition as a proxy measure): facilitated self-help versus treatment as usual

Three studies (N=1,136) found no evidence for clinically or statistically significant effects of facilitated self-help relative to treatment as usual on attrition (p=0.22) (Table 217).

Table 217: Summary of findings table for effects of facilitated self-help compared with treatment as usual on retention in services or treatment acceptability (using attrition as a proxy measure)

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Attrition: Facilitated self-help versus TAU				
Drop-out Incomplete data at endpoint Follow-up: 15-20 weeks	Study population		RR 0.94 (0.85 to 1.04)	1,136 (3 studies)	⊕⊕⊕⊕ high	
	577 per 1000	542 per 1000 (490 to 600)				
	Moderate					
	417 per 1000	392 per 1000 (354 to 434)				

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

Retention in services and treatment acceptability (using attrition as a proxy measure): listening visits versus treatment as usual

Three studies (N=1,211) found no evidence for clinically or statistically significant effects of listening visits relative to treatment as usual on attrition (p=0.15) (Table 218).

Table 218: Summary of findings table for effects of listening visits compared with treatment as usual on retention in services or treatment acceptability (using attrition as a proxy measure)

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Attrition: Listening visits versus TAU				
Drop-out Incomplete data at endpoint Follow-up: 20-52 weeks	Study population		RR 1.22 (0.93 to 1.6)	1,211 (3 studies)	⊕⊕⊖⊖ low1,2	
	131 per 1000	160 per 1000 (122 to 210)				
	Moderate					
	102 per 1000	124 per 1000 (95 to 163)				

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total number of events is less than 300 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Retention in services and treatment acceptability (using attrition as a proxy measure): directive counselling versus treatment as usual

A single study (N=146) found no evidence for clinically or statistically significant effects of directive counselling relative to treatment as usual on attrition (p=0.32) (Table 219).

Table 219: Summary of findings table for effects of directive counselling compared with treatment as usual on retention in services or treatment acceptability (using attrition as a proxy measure)

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Attrition: Directive counselling versus TAU				
Drop-out Incomplete data at endpoint Follow-up: mean 12 weeks	Study population		RR 0.8 (0.51 to 1.25)	146 (1 study)	⊕⊕⊖⊖ low1,2	
	455 per 1000	364 per 1000 (232 to 568)				
	Moderate					
	455 per 1000	364 per 1000 (232 to 569)				

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total number of events is less than 300 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Retention in services and treatment acceptability (using attrition as a proxy measure): post-miscarriage counselling versus treatment as usual or enhanced treatment as usual

Two studies (N=99) found no evidence for clinically or statistically significant effects of post-miscarriage counselling relative to treatment as usual or enhanced treatment as usual (medical investigations into causes of miscarriage without counselling) on attrition (p=0.63) (Table 220).

Table 220: Summary of findings table for effects of post-miscarriage counselling compared with treatment as usual or enhanced treatment as usual on retention in services or treatment acceptability (using attrition as a proxy measure)

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Attrition: post-miscarriage counselling versus TAU/Enhanced TAU				
Drop-out Incomplete data at endpoint Follow-up: 2-7 weeks	Study population		RR 0.81 (0.35 to 1.89)	99 (2 studies)	⊕⊕⊖⊖ low1,2	
	200 per 1000	162 per 1000 (70 to 378)				
	Moderate					
	209 per 1000	169 per 1000 (73 to 395)				

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total number of events is less than 300 (a threshold rule-of-thumb)
[2] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Retention in services and treatment acceptability (using attrition as a proxy measure): post-traumatic birth counselling versus treatment as usual

A single study (N=103) reported no drop-out from post-traumatic birth counselling or treatment as usual and it was therefore not possible to calculate an effect size (Table 221).

Table 221: Summary of findings table for effects of post-traumatic birth counselling compared with treatment as usual on retention in services or treatment acceptability (using attrition as a proxy measure)

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	 Corresponding risk Attrition: post-traumatic birth counselling versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Drop-out Incomplete data at endpoint Follow-up: mean 13 weeks	See comment	See comment	Not estimable	103 (1 study)	See comment	

*The basis for the **assumed risk** (for example, the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% CI) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

Retention in services and treatment acceptability (using attrition as a proxy measure): social support versus treatment as usual

Three studies (N=807) found evidence for a moderate effect of social support relative to treatment as usual on attrition with higher drop-out associated with peer-mediated support or a support group (p=0.18). However, this effect was not statistically significant due to very serious imprecision (Table 222).

Table 222: Summary of findings table for effects of social support compared with treatment as usual on retention in services or treatment acceptability (using attrition as a proxy measure)

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	 Corresponding risk Attrition: Social support versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Drop-out Incomplete data at endpoint Follow-up: 8-14 weeks	Study population		RR 1.49 (0.83 to 2.68)	807 (3 studies)	⊕⊕⊖⊖ low1,2	
	92 per 1000	136 per 1000 (76 to 245)				
	Moderate					
	46 per 1000	69 per 1000 (38 to 123)				

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total number of events is less than 300 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Retention in services and treatment acceptability (using attrition as a proxy measure): psychologically (CBT/IPT)-informed psychoeducation versus treatment as usual or enhanced treatment as usual

Thirteen studies (N=2,375) found no evidence for clinically or statistically significant effects of psychologically (CBT/IPT)-informed psychoeducational interventions relative to treatment as usual or enhanced treatment as usual on attrition (p=0.15) (Table 223).

Table 223: Summary of findings table for effects of psychologically (CBT/IPT)-informed psychoeducation compared with treatment as usual or enhanced

treatment as usual on retention in services or treatment acceptability (using attrition as a proxy measure)

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Attrition: Psychologically (CBT/IPT)-informed psychoeducation versus TAU/Enhanced TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Drop-out Incomplete data at endpoint Follow-up: 4-31 weeks	Study population 138 per 1000 Moderate 80 per 1000	161 per 1000 (130 to 200) 94 per 1000 (75 to 116)	RR 1.17 (0.94 to 1.45)	2,375 (13 studies)	⊕⊕⊕⊖ moderate1	

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Retention in services and treatment acceptability (using attrition as a proxy measure): non-mental health-focused education and support versus treatment as usual

A single study (N=331) found no evidence for a clinically or statistically significant effect of a non-mental health-focused education and support intervention relative to treatment as usual on attrition (p=0.73) (Table 224).

Table 224: Summary of findings table for effects of non-mental health-focused education and support compared with treatment as usual on retention in services or treatment acceptability (using attrition as a proxy measure)

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Attrition: Non-mental health-focused education and support versus TAU				
Drop-out Incomplete data at endpoint Follow-up: mean 12 weeks	Study population		RR 0.96 (0.75 to 1.22)	331 (1 study)	⊕⊕⊖⊖ low[1,2]	
	442 per 1000	424 per 1000 (331 to 539)				
	Moderate					
	442 per 1000	424 per 1000 (332 to 539)				

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total number of events is less than 300 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Retention in services and treatment acceptability (using attrition as a proxy measure): home visits versus treatment as usual

Four studies (N=1,252) found no evidence for clinically or statistically significant effects of home visits relative to treatment as usual on attrition (p=0.56) (Table 225).

Table 225: Summary of findings table for effects of home visits compared with treatment as usual on retention in services or treatment acceptability (using attrition as a proxy measure)

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Attrition: Home visits versus TAU				
Drop-out Incomplete data at endpoint Follow-up: 6-52 weeks	Study population		RR 1.07 (0.86 to 1.32)	1,252 (4 studies)	⊕⊕⊖⊖ low1,2	
	207 per 1000	221 per 1000 (178 to 273)				
	Moderate					
	196 per 1000	210 per 1000 (169 to 259)				

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total number of events is less than 300 (a threshold rule-of-thumb)
[2] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Retention in services and treatment acceptability (using attrition as a proxy measure): mother-infant relationship interventions versus treatment as usual or enhanced treatment as usual

Five studies (N=576) found no evidence for clinically or statistically significant effects of mother–infant relationship interventions relative to treatment as usual or enhanced treatment as usual on attrition (p=0.22) (Table 226).

Table 226: Summary of findings table for effects of mother–infant relationship interventions compared with treatment as usual or enhanced treatment as usual

on retention in services or treatment acceptability (using attrition as a proxy measure)

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Attrition: Mother-infant relationship interventions versus TAU/Enhanced TAU				
Drop-out Incomplete data at endpoint Follow-up: 5-28 weeks	Study population		RR 0.84 (0.63 to 1.12)	576 (5 studies)	⊕⊕⊖⊖ low1,2	
	238 per 1000	200 per 1000 (150 to 267)				
	Moderate					
	143 per 1000	120 per 1000 (90 to 160)				

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total number of events is less than 300 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Retention in services and treatment acceptability (using attrition as a proxy measure): mother-infant relationship intervention with video feedback versus mother–infant relationship intervention with verbal feedback

A single study (N=51) found no clinically or statistically significant difference on attrition (p=0.79) between a mother–infant relationship intervention with video feedback and a mother–infant relationship intervention with verbal feedback (Table 227).

Table 227: Summary of findings table for effects of mother–infant relationship intervention with video feedback compared with mother–infant relationship

intervention with verbal feedback on retention in services or treatment acceptability (using attrition as a proxy measure)

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Attrition: Mother-infant relationship intervention with video feedback versus mother–infant relationship intervention with verbal feedback				
Drop-out Incomplete data at endpoint Follow-up: mean 3 weeks	Study population		RR 0.87 (0.3 to 2.48)	51 (1 study)	⊕⊕⊖⊖ low1,2	
	231 per 1000	201 per 1000 (69 to 572)				
	Moderate					
	231 per 1000	201 per 1000 (69 to 573)				

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total number of events is less than 300 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Retention in services and treatment acceptability (using attrition as a proxy measure): mother-infant relationship intervention (and facilitated self-help) versus listening visits (and facilitated self-help)

There was single study (N=80) evidence for a moderate to large effect on attrition of a mother–infant relationship intervention relative to listening visits (in addition to facilitated self-help aimed at the eating disorder for both groups) with higher drop-out observed in the mother–infant relationship intervention group (p=0.56). However, this effect was not statistically significant due to very serious imprecision (Table 228).

Table 228: Summary of findings table for effects of mother–infant relationship intervention (and facilitated self-help) compared with listening visits (and

facilitated self-help) on retention in services or treatment acceptability (using attrition as a proxy measure)

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Attrition: Mother-infant relationship intervention (and facilitated self-help) versus listening visits (and facilitated self-help)	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Drop-out Incomplete data at endpoint Follow-up: mean 35 weeks	Study population		RR 2 (0.19 to 21.18)	80 (1 study)	⊕⊕⊖⊖ low1,2	
	25 per 1000	50 per 1000 (5 to 530)				
	Moderate					
	25 per 1000	50 per 1000 (5 to 530)				

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total number of events is less than 300 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Retention in services and treatment acceptability (using attrition as a proxy measure): co-parenting intervention versus enhanced treatment as usual

A single study (N=29) reported no drop-out from a co-parenting intervention or enhanced treatment as usual (monitoring) and it was therefore not possible to calculate an effect size (Table 229).

Table 229: Summary of findings table for effects of co-parenting intervention compared with enhanced treatment as usual on retention in services or treatment acceptability (using attrition as a proxy measure)

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	 Corresponding risk Attrition: Co-parenting intervention versus Enhanced TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Drop-out Incomplete data at endpoint Follow-up: mean 6 weeks	See comment	See comment	Not estimable	29 (1 study)	See comment	

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

Retention in services and treatment acceptability (using attrition as a proxy measure): music therapy during birth versus treatment as usual

A single study (N=141) found no evidence for a clinically or statistically significant effect of music therapy during birth relative to treatment as usual on attrition (p=0.61) (Table 230).

Table 230: Summary of findings table for effects of music therapy during birth compared with treatment as usual on retention in services or treatment acceptability (using attrition as a proxy measure)

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Attrition: Music therapy during birth versus TAU				
Drop-out Incomplete data at endpoint Follow-up: mean 3 weeks	Study population		RR 0.81 (0.36 to 1.83)	141 (1 study)	⊕⊕⊖⊖ low1,2	
	157 per 1000	127 per 1000 (57 to 288)				
	Moderate					
	157 per 1000	127 per 1000 (57 to 287)				

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total number of events is less than 300 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Retention in services and treatment acceptability (using attrition as a proxy measure): psychosomatic interventions versus treatment as usual

Two studies (N=276) found no evidence for clinically or statistically significant effects of psychosomatic interventions relative to treatment as usual on attrition (p=0.56) (Table 231).

Table 231: Summary of findings table for effects of psychosomatic interventions compared with treatment as usual on retention in services or treatment acceptability (using attrition as a proxy measure)

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	 Corresponding risk Attrition: Psychosomatic intervention versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Drop-out Incomplete data at endpoint Follow-up: 34-52 weeks	Study population		RR 0.87 (0.54 to 1.39)	276 (2 studies)	⊕⊖⊖⊖ very low1,2,3	
	413 per 1000	359 per 1000 (223 to 574)				
	Moderate					
	435 per 1000	378 per 1000 (235 to 605)				

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 There was evidence of moderate to substantial heterogeneity between effect sizes
2 Total number of events is less than 300 (a threshold rule-of-thumb)
3 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Retention in services and treatment acceptability (using attrition as a proxy measure): mindfulness training versus enhanced treatment as usual

A single study (N=47) found evidence for a moderate effect of mindfulness training relative to enhanced treatment as usual (non-mental health-focused education and support [book]) on attrition (p=0.73), with higher drop-out in the mindfulness training group. However, this effect was not statistically significant due to very serious imprecision (Table 232).

Table 232: Summary of findings table for effects of mindfulness training compared with enhanced treatment as usual on retention in services or treatment acceptability (using attrition as a proxy measure)

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Attrition: Mindfulness training versus Enhanced TAU				
Drop-out Incomplete data at endpoint Follow-up: mean 6 weeks	Study population		RR 1.28 (0.32 to 5.1)	47 (1 study)	⊕⊕⊖⊖ low1,2	
	130 per 1000	167 per 1000 (42 to 665)				
	Moderate					
	130 per 1000	166 per 1000 (42 to 663)				

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total number of events is less than 300 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

7.5.19 Clinical evidence for effects on infant service use (by intervention)

Summary of findings can be found in the tables presented in this section. The full GRADE evidence profiles and associated forest plots can be found in Appendix 22 and Appendix 19, respectively.

Infant service use: facilitated self-help versus treatment as usual

A single study (N=57-83) found evidence for moderate effects of facilitated self-help on reducing infant hospital use relative to treatment as usual (p=0.22-0.39). However, these effects were not statistically significant due to very serious imprecision and this study found no evidence for clinically or statistically significant effects of facilitated self-help on a continuous measure of infant hospital use (p=0.66) (Table 233).

Table 233: Summary of findings table for effects of facilitated self-help compared with treatment as usual on infant service use

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Infant service use: Facilitated self-help versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Infant hospital Post-Treatment (service utilisation at endpoint) – ITT analysis AD-SUS: Infant hospital Follow-up: mean 17 weeks	Study population 500 per 1000	365 per 1000 (220 to 605)	RR 0.73 (0.44 to 1.21)	83 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	
	Moderate 500 per 1000	365 per 1000 (220 to 605)				
Infant hospital Post-Treatment (service utilisation at endpoint) – available case analysis AD-SUS: Infant hospital Follow-up: mean 17 weeks	Study population 222 per 1000	133 per 1000 (42 to 422)	RR 0.6 (0.19 to 1.9)	57 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	
	Moderate 222 per 1000	133 per 1000 (42 to 422)				
Infant hospital Post-Treatment (service utilisation at endpoint) – available case analysis AD-SUS: Infant hospital Follow-up: mean 17 weeks		The mean infant hospital post-treatment (service utilisation at endpoint) – available case analysis in the intervention groups was 0.12 standard deviations lower (0.64 lower to 0.4 higher)		57 (1 study)	⊕⊖⊖⊖ very low[2,3,4]	SMD -0.12 (-0.64 to 0.4)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total number of events is less than 300 (a threshold rule-of-thumb)
[2] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
[3] Paper omits data
[4] Total population size is less than 400 (a threshold rule-of-thumb)

Infant service use: listening visits versus treatment as usual

There was single study (N=597-731) evidence for moderate effects of listening visits relative to treatment as usual on infant visits to an NHS health visitor at clinic at long-term follow-up with higher service usage in the listening visits group (p=0.06-0.15). However, these effects were not statistically significant due to very serious imprecision and the effects on this outcome measure were not clinically or statistically significant at post-treatment (p=0.81-0.95). This study also found evidence for a moderate effect of listening visits on visits for an infant from an NHS health visitor at home (with more visits observed for the intervention group) when using an available case analysis approach (p=0.08). However, again effect estimates were very imprecise and for this outcome measure the effect was not clinically or statistically significant when an ITT analysis approach was adopted (p=0.55). Moreover, it was unclear from the study whether this service usage was independent from the intervention, and thus, this outcome measure may be interpreted as a compliance measure. A moderate effect of listening visits relative to treatment as usual were observed on infant skin ointment usage with lower usage observed in the intervention group (p=0.006-0.01). A large effect of listening visits on infant asthma medication use was also observed (p=0.10) with lower usage in the listening visit relative to the treatment as usual group when an available case analysis approach was used. However, the effect estimate was very imprecise and the ITT analysis did not reveal any clinically or statistically significant effects on infant use of asthma medication (p=0.31). A small and statistically significant effect of listening visits on infant visits to the GP was found at post-treatment (p=0.02), however, this effect estimate did not meet criteria for clinical significance (as SMD<0.5) and effects were not clinically or statistically significant for infant visits to the GP at long-term follow-up (p=0.40-0.85). Finally, there was no evidence found for clinically or statistically significant effects of listening visits on infant use of hospital (p=0.61-0.75), infant visits to emergency department (measured at post-treatment [p=0.57-0.98] and long-term follow-up [p=0.51-0.87]), any infant medication use (p=0.27-0.47), or antibiotic use (p=0.95-0.96) (Table 234).

Table 234: Summary of findings table for effects of listening visits compared with treatment as usual on infant service use

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	 Corresponding risk Infant service use: Listening visits versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Infant hospital Post-Treatment (service utilisation at endpoint) – ITT analysis Child Health Service Use – Visits to hospital doctors (previous month) Follow-up: mean 52 weeks	Study population 237 per 1000 Moderate 237 per 1000	 218 per 1000 (159 to 299) 218 per 1000 (159 to 299)	RR 0.92 (0.67 to 1.26)	731 (1 study)	⊕⊖⊖⊖ very low1,2,3	
Infant hospital Post-Treatment (service utilisation at endpoint) – available case analysis Child Health Service Use – Visits to hospital doctors (previous month) Follow-up: mean 52 weeks	Study population 143 per 1000 Moderate 143 per 1000	 133 per 1000 (86 to 208) 133 per 1000 (86 to 207)	RR 0.93 (0.6 to 1.45)	653 (1 study)	⊕⊖⊖⊖ very low1,2,3	
Visit to emergency department Post-Treatment (service utilisation measured at endpoint) – ITT analysis Child Health Service Use – Visits to emergency department (previous month) Follow-up: mean 52 weeks	Study population 381 per 1000 Moderate 381 per 1000	 381 per 1000 (309 to 473) 381 per 1000 (309 to 472)	RR 1 (0.81 to 1.24)	731 (1 study)	⊕⊕⊖⊖ low1,3	
Visit to emergency department Post-Treatment (service utilisation measured at endpoint) – available case analysis Child Health Service Use – Visits to emergency department	Study population 266 per 1000 Moderate 266 per 1000	 290 per 1000 (218 to 386) 290 per 1000 (218 to 386)	RR 1.09 (0.82 to 1.45)	621 (1 study)	⊕⊖⊖⊖ very low1,2,3	

(previous month) Follow-up: mean 52 weeks				
Visit to NHS health visitor at clinic Post-Treatment (service utilisation [in past month] at endpoint) – ITT analysis Child Health Service Use – Visits to NHS health visitor at clinic (previous month) Follow-up: mean 52 weeks	Study population 392 per 1000 — 381 per 1000 (310 to 471) Moderate 392 per 1000 — 380 per 1000 (310 to 470)	RR 0.97 (0.79 to 1.2)	731 (1 study)	⊕⊕⊖⊖ low1,3
Visit to NHS health visitor at clinic Post-Treatment (service utilisation [in past month] at endpoint) – available case analysis Child Health Service Use – Visits to NHS health visitor at clinic (previous month) Follow-up: mean 52 weeks	Study population 318 per 1000 — 314 per 1000 (245 to 410) Moderate 318 per 1000 — 315 per 1000 (245 to 410)	RR 0.99 (0.77 to 1.29)	653 (1 study)	⊕⊕⊖⊖ low1,2,3
Visit from NHS health visitor at home Post-Treatment (service utilisation [in past month] at endpoint) – by intervention Child Health Service Use – Visits from NHS health visitor at home (previous month) Follow-up: mean 52 weeks	Study population 141 per 1000 — 159 per 1000 (107 to 235) Moderate 141 per 1000 — 159 per 1000 (107 to 235)	RR 1.13 (0.76 to 1.67)	731 (1 study)	⊕⊖⊖⊖ very low1,2,3
Visit from NHS health visitor at home Post-Treatment (service utilisation [in past month] at endpoint) – by intervention Child Health Service Use – Visits from NHS health visitor at home (previous month) Follow-up: mean 52 weeks	Study population 35 per 1000 — 67 per 1000 (32 to 139) Moderate 35 per 1000 — 67 per 1000 (32 to 140)	RR 1.91 (0.92 to 4)	653 (1 study)	⊕⊖⊖⊖ very low1,2,3
Visit to GP Post-Treatment (service utilisation [in past	Study population 546 per 1000 — 442 per 1000 (371 to 529)	RR 0.81 (0.68 to 0.97)	731 (1 study)	⊕⊕⊕⊖ moderate3

month] at endpoint) – ITT analysis Child Health Service Use – Visit to GP (previous month) Follow-up: mean 52 weeks	Moderate 546 per 1000 | 442 per 1000 (371 to 530)			
Visit to GP Post-Treatment (service utilisation [in past month] at endpoint) – available case analysis Child Health Service Use – Visit to GP (previous month) Follow-up: mean 52 weeks	Study population 490 per 1000 | 382 per 1000 (309 to 475) Moderate 490 per 1000 | 382 per 1000 (309 to 475)	RR 0.78 (0.63 to 0.97)	653 (1 study)	⨁⨁⨁⊖ moderate3
Any medication Post-Treatment (medication use [in past week] at endpoint) – ITT analysis Child medication use: any medication (previous week) Follow-up: mean 52 weeks	Study population 668 per 1000 | 708 per 1000 (634 to 795) Moderate 668 per 1000 | 708 per 1000 (635 to 795)	RR 1.06 (0.95 to 1.19)	731 (1 study)	⨁⨁⨁⊖ moderate3
Any medication Post-Treatment (past medication use measured at endpoint) – by intervention Child medication use: any medication (previous week) Follow-up: mean 52 weeks	Study population 630 per 1000 | 662 per 1000 (580 to 750) Moderate 630 per 1000 | 661 per 1000 (580 to 750)	RR 1.05 (0.92 to 1.19)	657 (1 study)	⨁⨁⨁⊖ moderate3
Antibiotics Post-Treatment (medication use [in past week] at endpoint) – ITT analysis Child medication use: antibiotics (previous week) Follow-up: mean 52 weeks	Study population 193 per 1000 | 191 per 1000 (135 to 269) Moderate 193 per 1000 | 191 per 1000 (135 to 268)	RR 0.99 (0.7 to 1.39)	731 (1 study)	⨁⊖⊖⊖ very low1,2,3
Antibiotics Post-Treatment (medication use [in past week] at endpoint) – available case analysis Child medication use: antibiotics (previous week)	Study population 102 per 1000 | 103 per 1000 (61 to 174) Moderate 102 per 1000 | 103 per 1000 (61 to 174)	RR 1.01 (0.6 to 1.71)	657 (1 study)	⨁⊖⊖⊖ very low1,2,3

Follow-up: mean 52 weeks				
Asthma medication Post-Treatment (medication use [in past week] at endpoint) – ITT analysis Child medication use: asthma medication (previous week) Follow-up: mean 52 weeks	Study population 139 per 1000 — 110 per 1000 (69 to 173) Moderate 139 per 1000 — 110 per 1000 (69 to 174)	RR 0.79 (0.5 to 1.25)	731 (1 study)	⊕⊖⊖⊖ very low1,2,3
Asthma medication Post-Treatment (medication use [in past week] at endpoint) – available case analysis Child medication use: asthma medication (previous week) Follow-up: mean 52 weeks	Study population 41 per 1000 — 12 per 1000 (3 to 51) Moderate 41 per 1000 — 12 per 1000 (3 to 52)	RR 0.3 (0.07 to 1.26)	657 (1 study)	⊕⊖⊖⊖ very low1,2,3
Skin ointment Post-Treatment (medication use [in past week] at endpoint) – ITT analysis Child medication use: Skin ointment (previous week) Follow-up: mean 52 weeks	Study population 325 per 1000 — 224 per 1000 (166 to 302) Moderate 325 per 1000 — 224 per 1000 (166 to 302)	RR 0.69 (0.51 to 0.93)	731 (1 study)	⊕⊕⊖⊖ low1,3
Skin ointment Post-Treatment (medication use [in past week] at endpoint) – available case analysis Child medication use: Skin ointment (previous week) Follow-up: mean 52 weeks	Study population 248 per 1000 — 139 per 1000 (92 to 211) Moderate 248 per 1000 — 139 per 1000 (92 to 211)	RR 0.56 (0.37 to 0.85)	657 (1 study)	⊕⊕⊖⊖ low1,3
Visit to emergency department Long follow-up (service utilisation [in past month] at >24 week follow-up) – ITT analysis Child Health Service Use – Visits to emergency department (previous month) Follow-up: mean 78 weeks	Study population 339 per 1000 — 367 per 1000 (292 to 458) Moderate 339 per 1000 — 366 per 1000 (292 to 458)	RR 1.08 (0.86 to 1.35)	731 (1 study)	⊕⊕⊖⊖ low1,2,3

Visit to emergency department Long follow-up (service utilisation [in past month] at >24 week follow-up) – available case analysis Child Health Service Use – Visits to emergency department (previous month) Follow-up: mean 78 weeks	Study population 201 per 1000 — 195 per 1000 (133 to 285) Moderate 201 per 1000 — 195 per 1000 (133 to 285)	RR 0.97 (0.66 to 1.42)	597 (1 study)	⨁⊖⊖⊖ very low1,2,3
Visit to NHS health visitor at clinic Long follow-up (service utilisation [in past month] at >24 week follow-up) – ITT analysis Child Health Service Use – Visits to NHS health visitor at clinic (previous month) Follow-up: mean 78 weeks	Study population 263 per 1000 — 334 per 1000 (260 to 428) Moderate 263 per 1000 — 334 per 1000 (260 to 429)	RR 1.27 (0.99 to 1.63)	731 (1 study)	⨁⨁⊖⊖ low1,2,3
Visit to NHS health visitor at clinic Long follow-up (service utilisation [in past month] at >24 week follow-up) – available case analysis Child Health Service Use – Visits to NHS health visitor at clinic (previous month) Follow-up: mean 78 weeks	Study population 114 per 1000 — 159 per 1000 (100 to 250) Moderate 114 per 1000 — 158 per 1000 (100 to 250)	RR 1.39 (0.88 to 2.19)	601 (1 study)	⨁⊖⊖⊖ very low1,2,3
Visit to GP Long follow-up (service utilisation [in past month] at >24 week follow-up) – ITT analysis Child Health Service Use – Visit to GP (previous month) Follow-up: mean 78 weeks	Study population 505 per 1000 — 495 per 1000 (420 to 586) Moderate 506 per 1000 — 496 per 1000 (420 to 587)	RR 0.98 (0.83 to 1.16)	731 (1 study)	⨁⨁⨁⊖ moderate3
Visit to GP Long follow-up (service utilisation [in past month] at >24 week follow-up) – available	Study population 406 per 1000 — 365 per 1000 (288 to 467) Moderate	RR 0.9 (0.71 to 1.15)	601 (1 study)	⨁⨁⊖⊖ low1,2,3

case analysis Child Health Service Use – Visit to GP (previous month) Follow-up: mean 78 weeks	406 per 1000	365 per 1000 (288 to 467)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total number of events is less than 300 (a threshold rule-of-thumb)
[2] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
[3] Paper omits data

Infant service use: home visits versus treatment as usual

A single study (N=268-364) found evidence for a moderate effect of home visits on infant hospitalisations with a lower number observed in the intervention group relative to the treatment as usual group (p=0.009) when an available case analysis approach was used. A small and statistically significant effect on infant hospitalisations was also observed for the ITT analysis, however, the effect estimate no longer met criteria for clinical significance (as RR>0.75). Confidence in these effect estimates was low due to risk of bias concerns (statistically significant group differences at baseline) and the rule-of-thumb threshold for optimal information size (300 events) was not met. This same study found no evidence for clinically or statistically significant effects of home visits on the number of children who were seen in an emergency department department (p=0.55-0.57). Another single study (N=138) found evidence for a moderate effect of home visits but this time in favour of the treatment as usual group with a higher administration of medication to the child without the advice of a medical practitioner in the home visit group (p=0.15). However, confidence in this effect estimate was very low due to risk of bias concerns (statistically significant group differences at baseline) and very serious imprecision (optimal information size threshold not reached and 95% CI includes both no effect and appreciable harm) (Table 235).

Table 235: Summary of findings table for effects of home visits compared with treatment as usual on infant service use

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	 Corresponding risk Infant service use: Home visits versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Infant hospital Post-Treatment (service utilisation at endpoint) – ITT analysis Medical record: Child hospitalisations Follow-up: mean 104 weeks	Study population 573 per 1000 Moderate 573 per 1000	 464 per 1000 (378 to 567) 464 per 1000 (378 to 567)	RR 0.81 (0.66 to 0.99)	364 (1 study)	⊕⊕⊖⊖ low1,2	
Infant hospital Post-Treatment (service utilisation at endpoint) – available case analysis Medical record: Child hospitalisations Follow-up: mean 104 weeks	Study population 423 per 1000 Moderate 423 per 1000	 267 per 1000 (191 to 377) 266 per 1000 (190 to 376)	RR 0.63 (0.45 to 0.89)	268 (1 study)	⊕⊕⊖⊖ low1,2	
Visit to emergency department Post-Treatment (service utilisation measured at endpoint) – ITT analysis Medical record: Child seen in emergency department Follow-up: mean 104 weeks	Study population 838 per 1000 Moderate 838 per 1000	 863 per 1000 (788 to 938) 863 per 1000 (788 to 939)	RR 1.03 (0.94 to 1.12)	364 (1 study)	⊕⊕⊕⊖ moderate1	
Visit to emergency department Post-Treatment (service utilisation measured at endpoint) – available case analysis Medical record: Child seen in emergency department Follow-up: mean 104 weeks	Study population 781 per 1000 Moderate 838 per 1000	 812 per 1000 (719 to 914) 872 per 1000 (771 to 980)	RR 1.04 (0.92 to 1.17)	268 (1 study)	⊕⊕⊕⊖ moderate1	
Any medication Post-Treatment (past medication use	Study population 114 per 1000	 206 per 1000 (93 to 459)	RR 1.8 (0.81 to 4.02)	138 (1 study)	⊕⊖⊖⊖ very low2,3,4	

measured at endpoint) – available case analysis Study-specific child health questionnaire: administration of medication to child without advice of medical practitioner Follow-up: mean 52 weeks	Moderate	
	114 per 1000	205 per 1000 (92 to 458)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 High risk of selection bias due to unclear allocation concealment and statistically significant baseline differences in poor psychological resources (37% intervention group versus 50% control) and in prenatal enrolment (41% intervention group and 53% control)
2 Total number of events is less than 300 (a threshold rule-of-thumb)
3 High risk of selection bias due to statistically significant baseline group differences in: parity (54% of intervention group primiparous versus 33% of control); identification as indigenous Australian (9% of intervention versus 2% of control); mental illness of partner (3% of intervention versus 14% of control); history of postnatal depression (11% of intervention versus 28% of control); physical domestic abuse (2% of intervention versus 10% of control); potential for child abuse (mean CAPI score in intervention was 123 versus 159 in control, and elevated CAPI score for 12% of intervention group versus 30% of control group)
4 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Infant service use: mother-infant relationship interventions versus treatment as usual or enhanced treatment as usual

A single study (N=95-121) found low quality evidence for moderate harms associated with a mother–infant relationship intervention relative to enhanced treatment as usual (non-mental health-focused education and support [booklet about infant care]) on infant hospitalisation (after discharge from NICU) and contact with specialized healthcare services with higher infant service use in the intervention group (p=0.15-0.39) when an available case analysis approach was used. However, effects on infant hospitalisation and contact with specialized healthcare services were not clinically or statistically significant when an ITT analysis approach was adopted (p=0.13-0.32). This study found no evidence for clinically or statistically significant effects on contact with developmental/rehabilitation specialist (p=0.59-0.69), use of any medication (p=0.13-0.15), surgery after discharge from NICU (p=0.55-0.86), or use of oxygen therapy (p=0.64-0.95) (Table 236).

Table 236: Summary of findings table for effects of mother–infant relationship interventions compared with treatment as usual or enhanced treatment as usual on infant service use

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Infant service use: Mother-infant relationship interventions versus TAU/Enhanced TAU				
Infant hospital Post-Treatment (service utilisation at endpoint) – ITT analysis Infant service use: Rehospitalized after discharge from NICU Follow-up: mean 25 weeks	Study population		RR 1.21 (0.83 to 1.77)	121 (1 study)	⊕⊕⊖⊖ low1,2	
	426 per 1000	516 per 1000 (354 to 754)				
	Moderate					
	426 per 1000	515 per 1000 (354 to 754)				
Infant hospital Post-Treatment (service utilisation at endpoint) – available case analysis Infant service use: Rehospitalized after discharge from NICU Follow-up: mean 25 weeks	Study population		RR 1.29 (0.72 to 2.31)	95 (1 study)	⊕⊕⊖⊖ low1,2	
	286 per 1000	369 per 1000 (206 to 660)				
	Moderate					
	286 per 1000	369 per 1000 (206 to 661)				
Contact with specialized healthcare services Post-Treatment (service utilisation at endpoint) – ITT analysis Infant service use: Contact with specialized health care services Follow-up: mean 25 weeks	Study population		RR 1.2 (0.95 to 1.52)	121 (1 study)	⊕⊕⊖⊖ low1,2	
	639 per 1000	767 per 1000 (607 to 972)				
	Moderate					
	639 per 1000	767 per 1000 (607 to 971)				
Contact with specialized healthcare services Post-Treatment (service utilisation at endpoint) – available case analysis Infant service use: Contact with specialized health care services Follow-up: mean 25 weeks	Study population		RR 1.26 (0.92 to 1.73)	95 (1 study)	⊕⊕⊖⊖ low1,2	
	551 per 1000	694 per 1000 (507 to 953)				
	Moderate					
	551 per 1000	694 per 1000 (507 to 953)				
	Study population					

Contact with developmental/rehabilitation specialist Post-Treatment (service utilisation at endpoint) – ITT analysis Infant service use: Contact with developmental/rehabilitation specialist Follow-up: mean 25 weeks	689 per 1000	737 per 1000 (585 to 923)			
	Moderate				
	689 per 1000	737 per 1000 (586 to 923)	RR 1.07 (0.85 to 1.34)	121 (1 study)	⊕⊕⊖⊖ low1,2
Contact with developmental/rehabilitation specialist Post-Treatment (service utilisation at endpoint) – available case analysis Infant service use: Contact with developmental/rehabilitation specialist Follow-up: mean 25 weeks	Study population		RR 1.07 (0.78 to 1.45)	95 (1 study)	⊕⊕⊖⊖ low1,2
	612 per 1000	655 per 1000 (478 to 888)			
	Moderate				
	612 per 1000	655 per 1000 (477 to 887)			
Any medication Post-Treatment (medication use [in past week] at endpoint) – ITT analysis Infant service use: Medication Follow-up: mean 25 weeks	Study population		RR 1.15 (0.96 to 1.38)	121 (1 study)	⊕⊕⊖⊖ low1,2
	738 per 1000	848 per 1000 (708 to 1000)			
	Moderate				
	738 per 1000	849 per 1000 (708 to 1000)			
Any medication Post-Treatment (past medication use measured at endpoint) – available case analysis Infant service use: Medication Follow-up: mean 25 weeks	Study population		RR 1.19 (0.94 to 1.52)	95 (1 study)	⊕⊕⊖⊖ low1,2
	673 per 1000	801 per 1000 (633 to 1000)			
	Moderate				
	674 per 1000	802 per 1000 (634 to 1000)			
Surgery Post-Treatment (service utilisation at endpoint) – ITT analysis Infant service use: Surgery after discharge from NICU Follow-up: mean 25 weeks	Study population		RR 0.86 (0.52 to 1.42)	109 (1 study)	⊕⊕⊖⊖ low1,2
	388 per 1000	333 per 1000 (202 to 551)			
	Moderate				
	388 per 1000	334 per 1000 (202 to 551)			
Surgery Post-Treatment (service utilisation at endpoint) – available case analysis Infant service use: Surgery after discharge from NICU Follow-up: mean 25 weeks	Study population		RR 0.91 (0.33 to 2.52)	95 (1 study)	⊕⊕⊖⊖ low1,2
	143 per 1000	130 per 1000 (47 to 360)			
	Moderate				
	143 per 1000	130 per 1000 (47 to 360)			
Oxygen therapy Post-Treatment (service utilisation at endpoint) – ITT analysis Infant service use: Oxygen	Study population		RR 1.16 (0.62 to 2.17)	121 (1 study)	⊕⊕⊖⊖ low1,2
	230 per 1000	266 per 1000 (142 to 498)			
	Moderate				

therapy Follow-up: mean 25 weeks	230 per 1000	267 per 1000 (143 to 499)			
Oxygen therapy Post-Treatment (service utilisation at endpoint) – available case analysis Infant service use: Oxygen therapy Follow-up: mean 25 weeks	Study population 41 per 1000	44 per 1000 (7 to 296)	RR 1.07 (0.16 to 7.25)	95 (1 study)	⊕⊕⊖⊖ low1,2
	Moderate 41 per 1000	44 per 1000 (7 to 297)			

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total number of events is less than 300 (a threshold rule-of-thumb)
[2] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

7.5.20 Clinical evidence for effects on infant physical health (by intervention)

Summary of findings can be found in the tables presented in this section. The full GRADE evidence profiles and associated forest plots can be found in Appendix 22 and Appendix 19, respectively.

Infant physical health: structured psychological interventions (CBT or IPT) versus treatment as usual or enhanced treatment as usual

A single study (N=705-903) found evidence for a moderate effect of CBT relative to enhanced treatment as usual (home visits) on the incidence of severe infant diarrhoea with a lower incidence in the intervention group when an available case analysis approach was used (p=0.003). The ITT analysis of this outcome measure also found a statistically significant effect (p=0.01) but the effect estimate no longer met criteria for clinical significance (as RR>0.75). This same study found no evidence for clinically or statistically significant effects of CBT on measures of infant weight (underweight [p=0.18-0.24] or weight-for-age [p=0.09]). With the exception of one statistically but not clinically significant effect estimate this study also found no evidence for clinically or statistically significant effects of CBT on measures of infant height (stunted height [p=0.09-0.28] or height-for-age [p=0.002]) (Table 237).

Table 237: Summary of findings table for effects of structured psychological interventions (CBT or IPT) compared with treatment as usual or enhanced treatment as usual on infant physical health

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Infant physical health: Structured psychological interventions (CBT or IPT) versus TAU/Enhanced TAU				
Underweight Post-treatment (underweight at endpoint or first measurement) – ITT analysis Child is considered underweight if growth is less than the anthropometric cutoff of -2 SD below the median Weight-for-age Z score and Height-for-age Z score of the National Center for Health Statistics/WHO international references Follow-up: mean 52 weeks	Study population		RR 0.95 (0.87 to 1.03)	903 (1 study)	⊕⊕⊕⊕ high	
	723 per 1000	687 per 1000 (629 to 744)				
	Moderate					
	723 per 1000	687 per 1000 (629 to 745)				
Underweight Post-treatment (underweight at endpoint or first measurement) – available case analysis Child is considered underweight if growth is less than the anthropometric cutoff of -2 SD below the median WAZ and HAZ scores of the National Center for Health Statistics/WHO international references Follow-up: mean 52 weeks	Study population		RR 0.92 (0.82 to 1.04)	705 (1 study)	⊕⊕⊕⊕ high	
	646 per 1000	595 per 1000 (530 to 672)				
	Moderate					
	646 per 1000	594 per 1000 (530 to 672)				

Weight-for-age Post-treatment (mean z score at endpoint or first measurement) – available case analysis Weight-for-age Z score Follow-up: mean 52 weeks		The mean weight-for-age post-treatment (mean z score at endpoint or first measurement) – available case analysis in the intervention groups was 0.13 standard deviations higher (0.02 lower to 0.28 higher)		705 (1 study)	⨁⨁⨁⨁ high	SMD 0.13 (-0.02 to 0.28)
Stunted height Post-treatment (short-for-age at endpoint or first measurement) – ITT analysis Child is considered stunted if growth is less than the anthropometric cutoff of -2 SD below the median WAZ and HAZ scores of the National Center for Health Statistics/WHO international references Follow-up: mean 52 weeks	Study population 400 per 1000 Moderate 400 per 1000	364 per 1000 (308 to 432) 364 per 1000 (308 to 432)	RR 0.91 (0.77 to 1.08)	903 (1 study)	⨁⨁⨁⨁ high	
Stunted height Post-treatment (short-for-age at endpoint or first measurement) – available case analysis Child is considered stunted if growth is less than the anthropometric cutoff of -2 SD below the median WAZ and HAZ scores of the National Center for Health Statistics/WHO international references Follow-up: mean 52 weeks	Study population 235 per 1000 Moderate 235 per 1000	183 per 1000 (136 to 244) 183 per 1000 (136 to 244)	RR 0.78 (0.58 to 1.04)	705 (1 study)	⨁⨁⊖⊖ low1,2	
Height-for-age Post-treatment (mean z score at endpoint or first measurement) – available case analysis Height-for-age Z score Follow-up: mean 52 weeks		The mean height-for-age post-treatment (mean z score at endpoint or first measurement) – available case analysis in the intervention groups was 0.24 standard		705 (1 study)	⨁⨁⨁⨁ high	SMD 0.24 (0.09 to 0.39)

		deviations higher (0.09 to 0.39 higher)			
Diarrhoea Post-treatment (=>1 diarrhoea episodes [in past 2 weeks] at endpoint or first measurement) – ITT analysis Diarrhoea was defined as ≥3 unformed stools passed in 24 hours, and a diarrhoeal episode was defined as being separated from another episode by at least 3 diarrhoea-free days Follow-up: mean 52 weeks	Study population 555 per 1000	471 per 1000 (416 to 538)	RR 0.85 (0.75 to 0.97)	903 (1 study)	⊕⊕⊕⊕ high
	Moderate 555 per 1000	472 per 1000 (416 to 538)			
Diarrhoea Post-treatment (=>1 diarrhoea episodes [in past 2 weeks] at endpoint or first measurement) – available case analysis Diarrhoea was defined as ≥3 unformed stools passed in 24 hours, and a diarrhoeal episode was defined as being separated from another episode by at least 3 diarrhoea-free days Follow-up: mean 52 weeks	Study population 432 per 1000	324 per 1000 (268 to 389)	RR 0.75 (0.62 to 0.9)	705 (1 study)	⊕⊕⊕⊕ high
	Moderate 432 per 1000	324 per 1000 (268 to 389)			

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total number of events is less than 300 (a threshold rule-of-thumb)
[2] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Infant physical health: IPT versus support group

A single study (N=44) found no evidence for clinically or statistically significant differences between IPT and a support group for gestational age (p=0.33) or birthweight (p=0.78) (Table 238).

Table 238: Summary of findings table for effects of IPT compared with support group on infant physical health

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Infant physical health: IPT versus support group	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Gestational age Post-treatment (mean score at endpoint or first measurement) – available case analysis Follow-up: mean 12 weeks		The mean gestational age post-treatment (mean score at endpoint or first measurement) – available case analysis in the intervention groups was 0.3 standard deviations lower (0.89 lower to 0.3 higher)		44 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	SMD -0.3 (-0.89 to 0.3)
Birth weight Post-treatment (mean score at endpoint or first measurement) – available case analysis Follow-up: mean 12 weeks		The mean birth weight post-treatment (mean score at endpoint or first measurement) – available case analysis in the intervention groups was 0.08 standard deviations lower (0.67 lower to 0.51 higher)		44 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	SMD -0.08 (-0.67 to 0.51)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] High risk of selection bias due to unclear allocation concealment and statistically significant baseline differences with the control group showing a higher SES score/lower income and higher depression

(CES-D) mean score
[2] Total population size is less than 400 (a threshold rule-of-thumb)
[3] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Infant physical health: listening visits versus treatment as usual

There was single study (N=650-731) low quality evidence for a moderate effect of listening visits relative to treatment as usual on maternal concerns about their child's health when using an available case analysis approach (p=0.07). However, the ITT analysis did not find a clinically or statistically significant effect (p=0.12) (Table 239).

Table 239: Summary of findings table for effects of listening visits compared with treatment as usual on infant physical health

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Infant physical health: Listening visits versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Ill health Post-treatment (maternal concerns about child health at endpoint or first measurement) – ITT analysis Child health and development concerns (maternal assessment): Child's health Follow-up: mean 52 weeks	Study population 394 per 1000 Moderate 394 per 1000	327 per 1000 (260 to 414) 327 per 1000 (260 to 414)	RR 0.83 (0.66 to 1.05)	731 (1 study)	⊕⊕⊖⊖ low[1,2,3]	
Ill health Post-treatment (maternal concerns about child health at endpoint or first measurement) – available case analysis Child health and development concerns (maternal assessment): Child's health Follow-up: mean 52 weeks	Study population 320 per 1000 Moderate 320 per 1000	240 per 1000 (179 to 326) 240 per 1000 (179 to 326)	RR 0.75 (0.56 to 1.02)	650 (1 study)	⊕⊕⊖⊖ low[1,2,3]	

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total number of events is less than 300 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
3 Paper omits data

Infant physical health: social support versus treatment as usual

A single study (N=23) found no evidence for a clinically or statistically significant effect of peer-mediated support (with mother–infant relationship intervention content) relative to a waitlist control on infant cortisol levels (p=0.52) (Table 240).

Table 240: Summary of findings table for effects of social support compared with treatment as usual on infant physical health

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Infant physical health: Social support versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Infant cortisol levels Post-treatment (mean score at endpoint or first measurement) – available case analysis Follow-up: mean 12 weeks		The mean infant cortisol levels post-treatment (mean score at endpoint or first measurement) – available case analysis in the intervention groups was 0.28 standard deviations higher (0.56 lower to 1.12 higher)		23 (1 study)	⊕⊕⊖⊖ low[1,2]	SMD 0.28 (-0.56 to 1.12)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the

estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total population size is less than 400 (a threshold rule-of-thumb)
[2] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Infant physical health: psychologically (CBT/IPT)-informed psychoeducation versus treatment as usual or enhanced treatment as usual

A single study (N=46-53) found no evidence for clinically or statistically significant effects of a CBT-informed psychoeducational intervention relative to treatment as usual on infant stress assessed by the mother using a visual analogue scale (p=0.40) or infant cortisol levels measured at post-treatment (p=0.32) or long-term follow-up (p=0.72) (Table 241).

Table 241: Summary of findings table for effects of psychologically (CBT/IPT)-informed psychoeducation compared with treatment as usual or enhanced treatment as usual on infant physical health

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Infant physical health: Psychologically (CBT/IPT)-informed psychoeducation versus TAU/Enhanced TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Infant stress Post-treatment (mean score at endpoint or first measurement) – available case analysis VAS: Infant stress Follow-up: mean 101 weeks		The mean infant stress post-treatment (mean score at endpoint or first measurement) – available case analysis in the intervention groups was 0.25 standard deviations higher (0.33 lower to 0.83 higher)		46 (1 study)	⊕⊖⊖⊖ very low[1,2,3,4]	SMD 0.25 (-0.33 to 0.83)
Infant cortisol levels Post-treatment (mean score at endpoint or first measurement) –		The mean infant cortisol levels post-treatment (mean score at endpoint or first		53 (1 study)	⊕⊖⊖⊖ very low[1,2,3,4]	SMD -0.27 (-0.82 to 0.27)

available case analysis Average (morning/evening) cortisol (log scores) Follow-up: mean 49 weeks	measurement) – available case analysis in the intervention groups was 0.27 standard deviations lower (0.82 lower to 0.27 higher)			
Infant cortisol levels Long follow-up (mean score at >24 week follow-up) – available case analysis Average (morning/evening) cortisol (log scores) Follow-up: mean 101 weeks	The mean infant cortisol levels long follow-up (mean score at >24 week follow-up) – available case analysis in the intervention groups was 0.11 standard deviations lower (0.69 lower to 0.47 higher)	46 (1 study)	⊕⊖⊖⊖ very low[1,2,3,4]	SMD -0.11 (-0.69 to 0.47)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] High risk of selection bias due to statistically significant baseline/mid-treatment difference in average maternal salivary cortisol levels (0.62 in intervention group and 0.75 in control group)
[2] Total population size is less than 400 (a threshold rule-of-thumb)
[3] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
[4] Paper omits data

Infant physical health: mother-infant relationship intervention (and facilitated self-help) versus listening visits (and facilitated self-help)

A single study (N=77) found no evidence for a clinically or statistically significant effect of a mother–infant relationship intervention relative to listening visits (both of which were in addition to facilitated self-help aimed at the eating disorder) on infant weight (p=0.61) (Table 242).

Table 242: Summary of findings table for effects of mother–infant relationship intervention (and facilitated self-help) compared with listening visits (and facilitated self-help) on infant physical health

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Infant physical health: Mother-infant relationship intervention (and facilitated self-help) versus listening visits (and facilitated self-help)	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Weight-for-age Post-treatment (mean z score at endpoint or first measurement) – available case analysis Weight-for-age Z score Follow-up: mean 35 weeks		The mean weight-for-age post-treatment (mean z score at endpoint or first measurement) – available case analysis in the intervention groups was 0.12 standard deviations lower (0.56 lower to 0.33 higher)		77 (1 study)	⊕⊕⊖⊖ low1,2	SMD -0.12 (-0.56 to 0.33)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total population size is less than 400 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

7.5.21 Clinical evidence for effects on infant physical development (by intervention)

Summary of findings can be found in the tables presented in this section. The full GRADE evidence profiles and associated forest plots can be found in Appendix 22 and Appendix 19, respectively.

Infant physical development: CBT versus listening visits

A single study (N=34) found no evidence for a clinically or statistically significant difference between CBT and listening visits on infant motor development (p=0.54) (Table 243).

Table 243: Summary of findings table for effects of CBT compared with listening visits on infant physical development

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Infant physical development: CBT versus listening visits	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Infant motor development Post-treatment (mean score at endpoint or first measurement) – available case analysis Bayley Scales of Infant Development-Psychomotor development index		The mean infant motor development post-treatment (mean score at endpoint or first measurement) – available case analysis in the intervention groups was 0.21 standard deviations higher (0.47 lower to 0.9 higher)		34 (1 study)	⊕⊕⊖⊖ low[1,2]	SMD 0.21 (-0.47 to 0.9)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total population size is less than 400 (a threshold rule-of-thumb)
[2] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Infant physical development: listening visits versus treatment as usual

A single study (N=591-731) found very low quality evidence for a moderate effect of listening visits relative to treatment as usual on infant eating habits when an

available case analysis was used (p=0.05). However, an ITT analysis of infant eating habits found no evidence for a clinically or statistically significant treatment effect (p=0.40). This study also found no evidence for clinically or statistically significant effects of listening visits on infant sleeping habits (p=0.54-0.68) (Table 244).

Table 244: Summary of findings table for effects of listening visits compared with treatment as usual on infant physical development

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Infant physical development: Listening visits versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Infant eating habits Post-treatment (maternal concerns at endpoint or first measurement) – ITT analysis Child health and development concerns (maternal assessment): Child's eating habits Follow-up: mean 78 weeks	Study population 369 per 1000 Moderate 369 per 1000	332 per 1000 (265 to 420) 332 per 1000 (266 to 421)	RR 0.9 (0.72 to 1.14)	731 (1 study)	⊕⊕⊖⊖ low1,2,3	
Infant eating habits Post-treatment (maternal concerns at endpoint or first measurement) – available case analysis Child health and development concerns (maternal assessment): Child's eating habits Follow-up: mean 78 weeks	Study population 228 per 1000 Moderate 228 per 1000	148 per 1000 (96 to 225) 148 per 1000 (96 to 226)	RR 0.65 (0.42 to 0.99)	591 (1 study)	⊕⊖⊖⊖ very low1,3	
Infant sleeping habits Post-treatment (maternal concerns at endpoint or first measurement) – ITT analysis Child health and development concerns (maternal assessment): Child's sleeping habits	Study population 290 per 1000 Moderate 290 per 1000	305 per 1000 (238 to 395) 304 per 1000 (238 to 394)	RR 1.05 (0.82 to 1.36)	731 (1 study)	⊕⊕⊖⊖ low1,2,3	

Follow-up: mean 78 weeks					
Infant sleep problems Post-treatment (maternal report at endpoint or first measurement) – available case analysis Child health and development concerns (maternal assessment): Child's sleeping habits Follow-up: mean 78 weeks	Study population		RR 0.85 (0.51 to 1.43)	591 (1 study)	⊕⊖⊖⊖ very low1,2,3
	132 per 1000	112 per 1000 (67 to 188)			
	Moderate				
	132 per 1000	112 per 1000 (67 to 189)			

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total number of events is less than 300 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
3 Paper omits data

Infant physical development: home visits versus treatment as usual

A single study (N=249-364) found very low quality evidence for a moderate effect of home visits relative to treatment as usual on reducing infant motor development impairment when an available case analysis approach was used (p=0.28). However, the ITT analysis did not find a clinically or statistically significant effect (p=0.19). Another study (N=138) found no evidence for clinically or statistically significant effects of home visits on infant feeding problems (p=0.25) or infant sleep problems (p=0.28) (Table 245).

Table 245: Summary of findings table for effects of home visits compared with treatment as usual on infant physical development

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Infant physical development: Home visits versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Infant motor development Post-treatment (below threshold at endpoint or first measurement) – ITT analysis Bayley Scales of Infant Development-Psychomotor development index<85 Follow-up: mean 104 weeks	Study population 470 per 1000 Moderate 470 per 1000	404 per 1000 (320 to 508) 404 per 1000 (320 to 508)	RR 0.86 (0.68 to 1.08)	364 (1 study)	⊕⊖⊖⊖ very low1,2,3	
Infant motor development Post-treatment (below threshold at endpoint or first measurement) – available case analysis Bayley Scales of Infant Development-Psychomotor development index<85 Follow-up: mean 104 weeks	Study population 203 per 1000 Moderate 203 per 1000	150 per 1000 (87 to 260) 150 per 1000 (87 to 260)	RR 0.74 (0.43 to 1.28)	249 (1 study)	⊕⊖⊖⊖ very low1,2,3	
Infant feeding problems Post-treatment (mean score at endpoint or first measurement) – available case analysis Study-specific child health questionnaire: Feeding problems		The mean infant feeding problems post-treatment (mean score at endpoint or first measurement) – available case analysis in the intervention groups was 0.2 standard deviations higher (0.14 lower to 0.53 higher)		138 (1 study)	⊕⊖⊖⊖ very low3,4,5	SMD 0.2 (-0.14 to 0.53)

Follow-up: mean 52 weeks				
Infant sleep problems Post-treatment (mean score at endpoint or first measurement) – available case analysis Study-specific child health questionnaire: Sleeping problems Follow-up: mean 52 weeks	The mean infant sleep problems post-treatment (mean score at endpoint or first measurement) – available case analysis in the intervention groups was 0.18 standard deviations higher (0.15 lower to 0.52 higher)	138 (1 study)	⊕⊖⊖⊖ very low3,4,5	SMD 0.18 (-0.15 to 0.52)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 High risk of selection bias due to unclear allocation concealment and statistically significant baseline differences in poor psychological resources (37% intervention group versus 50% control) and in prenatal enrolment (41% intervention group and 53% control)
2 Total number of events is less than 300 (a threshold rule-of-thumb)
3 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
4 High risk of selection bias due to statistically significant baseline group differences in: parity (54% of intervention group primiparous versus 33% of control); identification as indigenous Australian (9% of intervention versus 2% of control); mental illness of partner (3% of intervention versus 14% of control); history of postnatal depression (11% of intervention versus 28% of control); physical domestic abuse (2% of intervention versus 10% of control); potential for child abuse (mean CAPI score in intervention was 123 versus 159 in control, and elevated CAPI score for 12% of intervention group versus 30% of control group)
5 Total population size is less than 400 (a threshold rule-of-thumb)

Infant physical development: mother-infant relationship interventions versus treatment as usual or enhanced treatment as usual

A single study (N=96) found no evidence for a clinically or statistically significant effect of a mother–infant relationship intervention relative to enhanced treatment as usual (non-mental health-focused education and support [booklet about infant care]) on infant motor development (p=0.56) (Table 246).

Table 246: Summary of findings table for effects of mother–infant relationship interventions compared with treatment as usual or enhanced treatment as usual on infant physical development

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Infant physical development: Mother-infant relationship interventions versus TAU/Enhanced TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Infant motor development Post-treatment (mean score at endpoint or first measurement) – available case analysis Bayley Scales of Infant Development-Motor Follow-up: mean 25 weeks		The mean infant motor development post-treatment (mean score at endpoint or first measurement) – available case analysis in the intervention groups was 0.12 standard deviations lower (0.52 lower to 0.28 higher)		96 (1 study)	⊕⊕⊖⊖ low1,2	SMD -0.12 (-0.52 to 0.28)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total population size is less than 400 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Infant physical development: infant sleep training (controlled crying) versus treatment as usual

There was low to very low quality evidence from two studies (N=184-272) for moderate effects of infant sleep training (controlled crying) relative to treatment as usual on infant sleep problems at post-treatment (p=0.13) and at short-term follow-up (p=0.03). Although clinical and statistical significance was not maintained at long-term follow-up (p=0.34) (Table 247).

Table 247: Summary of findings table for effects of infant sleep training (controlled crying) compared with treatment as usual on infant physical development

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Infant physical development: Infant sleep training (controlled crying) versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Infant sleep problems Post-treatment (maternal report at endpoint or first measurement) – available case analysis Maternal report: Infant sleep problem – Treatment non-response (no further detail reported) Follow-up: 9-13 weeks	Study population		RR 0.55 (0.25 to 1.19)	189 (2 studies)	⊕⊖⊖⊖ very low1,2,3	
	677 per 1000	373 per 1000 (169 to 806)				
	Moderate					
	661 per 1000	364 per 1000 (165 to 787)				
Infant sleep problems Short follow-up (maternal report at 9-16 week follow-up) – available case analysis Maternal report: Infant sleep problem – Treatment non-response (no further detail reported) Follow-up: 17-22 weeks	Study population		RR 0.73 (0.55 to 0.97)	184 (2 studies)	⊕⊕⊖⊖ low2	
	591 per 1000	431 per 1000 (325 to 573)				
	Moderate					
	577 per 1000	421 per 1000 (317 to 560)				
Infant sleep problems Long follow-up (maternal report at >24 week follow-up) – available case analysis Maternal report: Infant sleep problem – Treatment non-response (no further detail re Follow-up: mean 74 weeks	Study population		RR 0.84 (0.58 to 1.21)	272 (1 study)	⊕⊕⊖⊖ low2,3	
	326 per 1000	273 per 1000 (189 to 394)				
	Moderate					
	326 per 1000	274 per 1000 (189 to 394)				

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 There was evidence of substantial to considerable heterogeneity between effect sizes
2 Total number of events is less than 300 (a threshold rule-of-thumb)
3 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

7.5.22 Clinical evidence for effects on infant cognitive development (by intervention)

Summary of findings can be found in the tables presented in this section. The full GRADE evidence profiles and associated forest plots can be found in Appendix 22 and Appendix 19, respectively.

Infant cognitive development: CBT versus listening visits

A single study (N=34) found no evidence for a statistically or clinically significant difference between CBT and listening visits on infant IQ (p=0.10) (Table 248).

Table 248: Summary of findings table for effects of CBT compared with listening visits on infant cognitive development

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Infant cognitive development: CBT versus listening visits				
Infant cognitive development Post-treatment (mean score at endpoint or first measurement) – available case analysis Bayley Scales of Infant		The mean infant cognitive development post-treatment (mean score at endpoint or first measurement) – available case analysis in the intervention groups was 0.59 standard		34 (1 study)	⊕⊕⊖⊖ low1,2	SMD 0.59 (-0.11 to 1.29)

Development-Mental development index		deviations higher (0.11 lower to 1.29 higher)	

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total population size is less than 400 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Infant cognitive development: listening visits versus treatment as usual

A single study (N=591) found very low quality evidence for a large effect of listening visits relative to treatment as usual on maternal concerns about infant verbal development when an available case analysis approach was used (p=0.01). However, the ITT analysis for this outcome measure (N=731) was not clinically or statistically significant (p=0.37). This same study (N=640-731) also found no evidence for clinically or statistically significant effects of listening visits on maternal concerns about infant development (p=0.73-0.95) (Table 249).

Table 249: Summary of findings table for effects of listening visits compared with treatment as usual on infant cognitive development

Outcomes	**Illustrative comparative risks* (95% CI)**		**Relative effect (95% CI)**	**No. of Participants (studies)**	**Quality of the evidence (GRADE)**	**Comments**
	Assumed risk	**Corresponding risk**				
	Control	**Infant cognitive development: Listening visits versus TAU**				
Infant cognitive development Post-treatment (maternal concerns/below threshold at endpoint or first measurement) – ITT analysis Child health and development concerns (maternal assessment):	Study population		RR 0.93 (0.64 to 1.37)	731 (1 study)	⊕⊖⊖⊖ very low1,2,3	
	170 per 1000	158 per 1000 (109 to 232)				
	Moderate					
	170 per 1000	158 per 1000 (109 to 233)				

Child's development Follow-up: mean 52 weeks				
Infant cognitive development Post-treatment (maternal concerns/below threshold at endpoint or first measurement) – available case analysis Child health and development concerns (maternal assessment): Child's development Follow-up: mean 52 weeks	Study population 48 per 1000 — 50 per 1000 (23 to 108) Moderate 48 per 1000 — 49 per 1000 (23 to 108)	RR 1.03 (0.47 to 2.25)	640 (1 study)	⊕⊖⊖⊖ very low1,2,3
Infant verbal development Post-treatment (maternal concerns at endpoint or first measurement) – ITT analysis Child health and development concerns (maternal assessment): Child's speech Follow-up: mean 78 weeks	Study population 303 per 1000 — 267 per 1000 (203 to 351) Moderate 303 per 1000 — 267 per 1000 (203 to 351)	RR 0.88 (0.67 to 1.16)	731 (1 study)	⊕⊖⊖⊖ very low1,2,3
Infant verbal development Post-treatment (maternal concerns at endpoint or first measurement) – available case analysis Child health and development concerns (maternal assessment): Child's speech Follow-up: mean 78 weeks	Study population 147 per 1000 — 63 per 1000 (32 to 124) Moderate 147 per 1000 — 63 per 1000 (32 to 123)	RR 0.43 (0.22 to 0.84)	591 (1 study)	⊕⊖⊖⊖ very low1,3

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total number of events is less than 300 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
3 Paper omits data

Infant cognitive development: social support versus treatment as usual

A single study (N=48) found no evidence for a clinically or statistically significant effect of peer-mediated support (with mother–infant relationship intervention content) relative to a waitlist control on infant IQ (p=0.47) (Table 250).

Table 250: Summary of findings table for effects of social support compared with treatment as usual on infant cognitive development

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Infant cognitive development: Social support versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Infant cognitive development Post-treatment (mean score at endpoint or first measurement) – available case analysis Bayley Scales of Infant Development-Mental development index Follow-up: mean 12 weeks		The mean infant cognitive development post-treatment (mean score at endpoint or first measurement) – available case analysis in the intervention groups was 0.21 standard deviations lower (0.78 lower to 0.36 higher)		48 (1 study)	⊕⊕⊖⊖ low1,2	SMD -0.21 (-0.78 to 0.36)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total population size is less than 400 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Infant cognitive development: home visits versus treatment as usual

A single study (N=249-364) found no evidence for clinically or statistically significant effects of home visits relative to treatment as usual on infant intellectual impairment (p=0.08-0.12) (Table 251).

Table 251: Summary of findings table for effects of home visits compared with treatment as usual on infant cognitive development

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Infant cognitive development: Home visits versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Infant cognitive development Post-treatment (maternal concerns/below threshold at endpoint or first measurement) – ITT analysis Bayley Scales of Infant Development- Mental development index<85 Follow-up: mean 104 weeks	Study population 681 per 1000	593 per 1000 (504 to 695)	RR 0.87 (0.74 to 1.02)	364 (1 study)	⊕⊕⊖⊖ low1,2,3	
	Moderate 681 per 1000	592 per 1000 (504 to 695)				
Infant cognitive development Post-treatment (maternal concerns/below threshold at endpoint or first measurement) – available case analysis Bayley Scales of Infant Development- Mental development index<85 Follow-up: mean 104 weeks	Study population 520 per 1000	421 per 1000 (323 to 546)	RR 0.81 (0.62 to 1.05)	249 (1 study)	⊕⊖⊖⊖ very low1,2,3	
	Moderate 520 per 1000	421 per 1000 (322 to 546)				

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the

estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 High risk of selection bias due to unclear allocation concealment and statistically significant baseline differences in poor psychological resources (37% intervention group versus 50% control) and in prenatal enrolment (41% intervention group and 53% control)
2 Total number of events is less than 300 (a threshold rule-of-thumb)
3 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Infant cognitive development: mother-infant relationship interventions versus treatment as usual or enhanced treatment as usual

A single study (N=96) found no evidence of a clinically or statistically significant effect of a mother–infant relationship intervention relative to enhanced treatment as usual (non-mental health-focused education and support [booklet about infant care]) on infant IQ (p=0.74) and two studies (N=154) found no evidence for clinically or statistically significant effects of mother–infant relationship interventions relative to enhanced treatment as usual (non-mental health-focused education and support [booklet about infant care] or telephone support) on infant verbal development (p=0.58) (Table 252).

Table 252: Summary of findings table for effects of mother–infant relationship interventions compared with treatment as usual or enhanced treatment as usual on infant cognitive development

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Infant cognitive development: Mother-infant relationship interventions versus TAU/Enhanced TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Infant cognitive development Post-treatment (mean score at endpoint or first measurement) – available case analysis Bayley Scales of Infant Development-Cognitive Follow-up: mean 25 weeks		The mean infant cognitive development post-treatment (mean score at endpoint or first measurement) – available case analysis in the intervention groups was **0.07 standard deviations higher** (0.33 lower to 0.47 higher)		96 (1 study)	⊕⊕⊖⊖ **low**[1]	SMD 0.07 (-0.33 to 0.47)
Infant verbal development Post-treatment (mean score at endpoint or first measurement) –		The mean infant verbal development post-treatment (mean score at endpoint or first measurement) – available case analysis in the intervention		154 (2 studies)	⊕⊕⊖⊖ **low**[1]	SMD 0.1 (-0.25 to 0.45)

available case analysis Peabody Picture Vocabulary Test-Revisedor Bayley Scales of Infant Development-Language Follow-up: 25-271 weeks	groups was **0.1 standard deviations higher** (0.25 lower to 0.45 higher)

*The basis for the **assumed risk** (for example, the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% CI) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total population size is less than 400 (a threshold rule-of-thumb)

7.5.23 Clinical evidence for effects on infant emotional development (by intervention)

Summary of findings can be found in the tables presented in this section. The full GRADE evidence profiles and associated forest plots can be found in Appendix 22 and Appendix 19, respectively.

Infant emotional development: social support versus treatment as usual

A single study (N=51) found no evidence for a clinically or statistically significant effect of peer-mediated support (with mother–infant relationship intervention content) relative to waitlist control on maternal-rated infant 'difficult' temperament (p=0.25) (Table 253).

Table 253: Summary of findings table for effects of social support compared with treatment as usual on infant emotional development

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk Control	Corresponding risk Infant emotional development: Social support versus TAU				
Infant 'difficult' temperament Post-treatment (maternal-rated mean score at		The mean infant 'difficult' temperament post-treatment (maternal-rated mean		51 (1 study)	⊕⊕⊖⊖ low1,2	SMD 0.33 (-0.23 to 0.88)

endpoint or first measurement) – available case analysis Infant Characteristics Questionnaire Follow-up: mean 12 weeks	score at endpoint or first measurement) – available case analysis in the intervention groups was 0.33 standard deviations higher (0.23 lower to 0.88 higher)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total population size is less than 400 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Infant emotional development: home visits versus treatment as usual

There was single study (N=249) very low quality evidence for a moderate effect of home visits relative to treatment as usual on infant internalizing using an available case analysis approach (p=0.08). However, ITT analysis for this outcome measure (N=364) found no evidence for a clinically or statistically significant effect (p=0.08). This study (N=249-364) also found no evidence for clinically or statistically significant effects of home visits on infant externalizing (p=0.24-0.38). Another study (N=160-440) found a similar pattern of treatment effects on infant social withdrawal with low quality evidence for a moderate effect on a dichotomous measure using available case analysis (p=0.09) but no evidence for clinically or statistically significant effects on ITT analysis of the same dichotomous measure (p=0.25) or on a continuous measure of infant social withdrawal (p=1.00) (Table 254).

Table 254: Summary of findings table for effects of home visits compared with treatment as usual on infant emotional development

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Infant emotional development: Home visits versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Infant externalizing Post-treatment (symptomatology – above threshold at endpoint or first measurement) – ITT analysis Child Behaviour Checklist (CBCL/1.5-5): Externalising Follow-up: mean 104 weeks	Study population		RR 0.87 (0.7 to 1.09)	364 (1 study)	⊕⊖⊖⊖ very low1,2,3	
	486 per 1000	423 per 1000 (341 to 530)				
	Moderate					
	487 per 1000	424 per 1000 (341 to 531)				
Infant externalizing Post-treatment (symptomatology – above threshold at endpoint or first measurement) – available case analysis CBCL/1.5-5: Externalising Follow-up: mean 104 weeks	Study population		RR 0.8 (0.49 to 1.31)	249 (1 study)	⊕⊖⊖⊖ very low1,2,3	
	228 per 1000	182 per 1000 (112 to 298)				
	Moderate					
	228 per 1000	182 per 1000 (112 to 299)				
Infant internalizing Post-treatment (symptomatology – above threshold at endpoint or first measurement) – ITT analysis CBCL/1.5-5: Internalising Follow-up: mean 104 weeks	Study population		RR 0.81 (0.64 to 1.03)	364 (1 study)	⊕⊖⊖⊖ very low1,2,3	
	476 per 1000	385 per 1000 (304 to 490)				
	Moderate					
	476 per 1000	386 per 1000 (305 to 490)				
Infant internalizing Post-treatment (symptomatology – above threshold at	Study population		RR 0.6 (0.34 to 1.06)	249 (1 study)	⊕⊖⊖⊖ very low1,2,3	
	211 per 1000	127 per 1000 (72 to 224)				
	Moderate					

endpoint or first measurement) – available case analysis CBCL/1.5-5: Internalising Follow-up: mean 104 weeks	211 per 1000	127 per 1000 (72 to 224)				
Infant social withdrawal Post-treatment (symptomatology – above threshold at endpoint or first measurement) – ITT analysis Alarm Distress Baby Scale =>5 Follow-up: mean 87 weeks	Study population		RR 0.86 (0.66 to 1.12)	440 (1 study)	⊕⊕⊖⊖ low2,3	
	362 per 1000	312 per 1000 (239 to 406)				
	Moderate					
	362 per 1000	311 per 1000 (239 to 405)				
Infant social withdrawal Post-treatment (symptomatology – above threshold at endpoint or first measurement) – available case analysis Alarm Distress Baby Scale =>5 Follow-up: mean 87 weeks	Study population		RR 0.7 (0.46 to 1.06)	367 (1 study)	⊕⊕⊖⊖ low2,3	
	240 per 1000	168 per 1000 (111 to 255)				
	Moderate					
	240 per 1000	168 per 1000 (110 to 254)				
Infant social withdrawal Post-treatment (mean score at endpoint or first measurement) – available case analysis Alarm Distress Baby Scale Follow-up: mean 87 weeks		The mean infant social withdrawal post-treatment (mean score at endpoint or first measurement) – available case analysis in the intervention groups was 0 standard deviations higher (0.31 lower to 0.31 higher)		160 (1 study)	⊕⊕⊖⊖ low4	SMD 0 (-0.31 to 0.31)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] High risk of selection bias due to unclear allocation concealment and statistically significant baseline differences in poor psychological resources (37% intervention group versus 50% control) and in prenatal enrolment (41% intervention group and 53% control)
[2] Total number of events is less than 300 (a threshold rule-of-thumb)
[3] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
[4] Total population size is less than 400 (a threshold rule-of-thumb)

Infant emotional development: mother-infant relationship interventions versus treatment as usual or enhanced treatment as usual

Two studies (N=146) found no evidence for clinically or statistically significant effects of mother–infant relationship interventions relative to treatment as usual or enhanced treatment as usual on a continuous measure of infant adaptive behaviour (p=0.61). In addition, one of those studies (N=75-80) also found no evidence for clinically or statistically significant effects of mother–infant psychotherapy relative to treatment as usual on dichotomous measures of infant adaptive behaviour (p=0.58-0.62) (Table 255).

A single study (N=58-71) found no evidence for clinically or statistically significant effects of a mother–infant relationship intervention relative to enhanced treatment as usual (non-mental health-focused education and support [booklet about infant care]) on infant externalizing (p=0.72) or infant dysregulation (p=0.75) at post-treatment or infant externalizing at very long-term follow-up (p=0.60). The same study also found no clinically or statistically significant treatment effects on infant internalizing at post-treatment (p=0.21). However, at very long-term follow-up there was evidence for a large harm associated with a mother–infant relationship intervention with more severe infant internalizing mean scores observed in the intervention group relative to the enhanced treatment as usual group (p <0.00001). This study did, however, find low quality evidence for a large benefit of a mother–infant relationship intervention on infant self-esteem (p <0.00001) (Table 255).

Table 255: Summary of findings table for effects of mother–infant relationship interventions compared with treatment as usual or enhanced treatment as usual on infant emotional development

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Infant emotional development: Mother-infant relationship interventions versus TAU/Enhanced TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Infant adaptive behaviour Post-treatment (treatment response at endpoint or first measurement) – ITT analysis Ages and Stages Questionnaire: Social-Emotional (ASQ:SE): Treatment response (improvement-reliable change index) Follow-up: mean 26 weeks	Study population 175 per 1000 Moderate 175 per 1000	226 per 1000 (93 to 546) 226 per 1000 (93 to 546)	RR 1.29 (0.53 to 3.12)	80 (1 study)	⊕⊖⊖⊖ very low1,2,3	
Infant adaptive behaviour Post-treatment (treatment response at endpoint or first measurement) – available case analysis (ASQ:SE: Treatment response (improvement-reliable change index) Follow-up: mean 26 weeks	Study population 189 per 1000 Moderate 189 per 1000	236 per 1000 (98 to 569) 236 per 1000 (98 to 569)	RR 1.25 (0.52 to 3.01)	75 (1 study)	⊕⊖⊖⊖ very low1,2,3	
Infant adaptive behaviour Post-treatment (mean score at endpoint or first measurement) –		The mean infant adaptive behaviour post-treatment (mean score at endpoint or first measurement) –		146 (2 studies)	⊕⊖⊖⊖ very low1,3,4,5	SMD 0.21 (-0.59 to 1)

available case analysis ASQ:SE or Infant Toddler Social and Emotional Assessment: Competence Follow-up: 26-57 weeks	available case analysis in the intervention groups was 0.21 standard deviations higher (0.59 lower to 1 higher)			
Infant externalizing Post-treatment (mean score at endpoint or first measurement) – available case analysis Infant Toddler Social and Emotional Assessment: Externalizing Follow-up: mean 57 weeks	The mean infant externalizing post-treatment (mean score at endpoint or first measurement) – available case analysis in the intervention groups was 0.09 standard deviations higher (0.38 lower to 0.55 higher)	71 (1 study)	⨁⨁⊖⊖ low3,5	SMD 0.09 (-0.38 to 0.55)
Infant internalizing Post-treatment (mean score at endpoint or first measurement) – available case analysis Infant Toddler Social and Emotional Assessment: Internalizing Follow-up: mean 57 weeks	The mean infant internalizing post-treatment (mean score at endpoint or first measurement) – available case analysis in the intervention groups was 0.3 standard deviations higher (0.17 lower to 0.77 higher)	71 (1 study)	⨁⨁⊖⊖ low3,5	SMD 0.3 (-0.17 to 0.77)
Infant dysregulation Post-treatment (mean score at endpoint or first measurement) – available case analysis Infant Toddler Social and Emotional Assessment: Dysregulation Follow-up: mean 57 weeks	The mean infant dysregulation post-treatment (mean score at endpoint or first measurement) – available case analysis in the intervention groups was 0.08 standard deviations lower (0.54 lower to 0.39 higher)	71 (1 study)	⨁⨁⊖⊖ low3,5	SMD -0.08 (-0.54 to 0.39)
Infant self-esteem Post-treatment (mean score at endpoint or first measurement) – available case analysis	The mean infant self-esteem post-treatment (mean score at endpoint or first measurement) – available case analysis in the intervention	58 (1 study)	⨁⨁⊖⊖ low5	SMD 1.46 (0.88 to 2.05)

Puppet Interview: Child self-esteem Follow-up: mean 271 weeks	groups was 1.46 standard deviations higher (0.88 to 2.05 higher)			
Infant externalizing Very long Follow-up (mean score at >104 week follow-up) – available case analysis CBCL/1.5-5: Externalising Follow-up: mean 271 weeks	The mean infant externalizing very long follow-up (mean score at >104 week follow-up) – available case analysis in the intervention groups was 0.14 standard deviations lower (0.65 lower to 0.38 higher)	58 (1 study)	⨁⨁⨀⨀ low3,5	SMD -0.14 (-0.65 to 0.38)
Infant internalizing Very long Follow-up (mean score at >104 week follow-up) – available case analysis CBCL/1.5-5: Internalising Follow-up: mean 271 weeks	The mean infant internalizing very long follow-up (mean score at >104 week follow-up) – available case analysis in the intervention groups was 1.79 standard deviations higher (1.17 to 2.4 higher)	58 (1 study)	⨁⨁⨀⨀ low5	SMD 1.79 (1.17 to 2.4)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 High risk of selection bias due to statistically significant baseline difference in the age of infants (4.4 months old in intervention group versus 5.9 months old in TAU group)
2 Total number of events is less than 300 (a threshold rule-of-thumb)
3 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
4 There was evidence of substantial to considerable heterogeneity between effect sizes
5 Total population size is less than 400 (a threshold rule-of-thumb)

Infant emotional development: infant sleep training (controlled crying) versus treatment as usual

A single study (N=268) found no evidence for clinically or statistically significant effects of infant sleep training (controlled crying) on infant externalizing (p=0.60) or internalizing (p=0.86) (Table 256).

Table 256: Summary of findings table for effects of infant sleep training (controlled crying) compared with treatment as usual on infant emotional development

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Infant emotional development: Infant sleep training (controlled crying) versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Infant externalizing Post-treatment (mean score at endpoint or first measurement) – available case analysis CBCL - Externalising Follow-up: mean 74 weeks		The mean infant externalizing post-treatment (mean score at endpoint or first measurement) – available case analysis in the intervention groups was 0.07 standard deviations higher (0.17 lower to 0.31 higher)		268 (1 study)	⊕⊕⊕⊖ moderate[1]	SMD 0.07 (-0.17 to 0.31)
Infant internalizing Post-treatment (mean score at endpoint or first measurement) – available case analysis CBCL - Internalising Follow-up: mean 74 weeks		The mean infant internalizing post-treatment (mean score at endpoint or first measurement) – available case analysis in the intervention groups was 0.02 standard deviations higher (0.22 lower to 0.26 higher)		268 (1 study)	⊕⊕⊕⊖ moderate[1]	SMD 0.02 (-0.22 to 0.26)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total population size is less than 400 (a threshold rule-of-thumb)

7.5.24 Clinical evidence for prevention of neglect or abuse of the infant (by intervention)

Summary of findings can be found in the tables presented in this section. The full GRADE evidence profiles and associated forest plots can be found in Appendix 22 and Appendix 19, respectively.

Prevention of neglect or abuse of the infant: listening visits versus treatment as usual

A single study (N=596-731) found no evidence for clinically or statistically significant effects of listening visits relative to treatment as usual on the incidence of child injury requiring medical attention at post-treatment (p=0.78-0.97) or long-term follow-up (p=0.19-0.76) (Table 257).

Table 257: Summary of findings table for effects of listening visits compared with treatment as usual for prevention of neglect or abuse of the infant

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Prevention of neglect or abuse of the infant: Listening visits versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Child injury Post-treatment (Injury requiring medical attention at endpoint or first measurement) – ITT analysis Child Health Service Use – Injury requiring medical attention Follow-up: mean 52 weeks	Study population 234 per 1000 Moderate 234 per 1000	 236 per 1000 (173 to 318) 236 per 1000 (173 to 318)	RR 1.01 (0.74 to 1.36)	731 (1 study)	⊕⊕⊖⊖ low[1,2,3]	
Child injury Post-treatment (Injury requiring medical attention at endpoint or first measurement) – available case analysis Child Health Service Use – Injury requiring medical attention	Study population 138 per 1000 Moderate 138 per 1000	 146 per 1000 (95 to 226) 146 per 1000 (95 to 226)	RR 1.06 (0.69 to 1.64)	651 (1 study)	⊕⊖⊖⊖ very low[1,2,3]	

Follow-up: mean 52 weeks				
Child injury Long follow-up (Injury requiring medical attention at >24 week follow-up) – ITT analysis Child Health Service Use – Injury requiring medical attention Follow-up: mean 78 weeks	Study population 252 per 1000: 300 per 1000 (232 to 390) Moderate 252 per 1000: 300 per 1000 (232 to 391)	RR 1.19 (0.92 to 1.55)	731 (1 study)	⊕⊕⊖⊖ low[1,2,3]
Child injury Long follow-up (Injury requiring medical attention at >24 week follow-up) – by intervention Child Health Service Use – Injury requiring medical attention Follow-up: mean 78 weeks	Study population 91 per 1000: 83 per 1000 (45 to 153) Moderate 91 per 1000: 83 per 1000 (45 to 153)	RR 0.91 (0.49 to 1.68)	596 (1 study)	⊕⊖⊖⊖ very low[1,2,3]

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Total number of events is less than 300 (a threshold rule-of-thumb)
[2] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
[3] Paper omits data

Prevention of neglect or abuse of the infant: home visits versus treatment as usual

A single study (N=138) found evidence for a large effect of home visits relative to treatment as usual on preventing the child ingesting poison (p=0.14). However, confidence in this effect estimate was very low due to a high risk of selection bias (statistically significant group differences at baseline) and very serious imprecision. Single study analyses of the data from this and one other study found no evidence for clinically or statistically significant effects of home visits relative to treatment as usual on child injury (p=0.58-0.75), child protective service reports of all types (p=0.73-0.82), child protective service reports of neglect (p=0.71-0.78), or maternal use of punishment (p=0.50-0.68). There was also no evidence for a clinically

significant effect (although the effect was statistically significant) of home visits on a continuous measure of potential for child abuse (p=0.05) (Table 258).

Table 258: Summary of findings table for effects of home visits compared with treatment as usual for prevention of neglect or abuse of the infant

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Prevention of neglect or abuse of the infant: Home visits versus TAU	Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
Child injury Post-treatment (Injury requiring medical attention at endpoint or first measurement) – ITT analysis Medical record: Child injuries requiring medical care Follow-up: mean 104 weeks	Study population		RR 0.97 (0.78 to 1.19)	364 (1 study)	⊕⊕⊖⊖ low1,2	
	497 per 1000	482 per 1000 (388 to 592)				
	Moderate					
	497 per 1000	482 per 1000 (388 to 591)				
Child injury Post-treatment (Injury requiring medical attention at endpoint or first measurement) – available case analysis Medical record: Child injuries requiring medical care Follow-up: mean 104 weeks	Study population		RR 0.9 (0.63 to 1.3)	268 (1 study)	⊕⊖⊖⊖ very low1,2,3	
	321 per 1000	289 per 1000 (202 to 418)				
	Moderate					
	321 per 1000	289 per 1000 (202 to 417)				
Ingestion of poison Post-treatment (incidence during trial measured at endpoint or first measurement) – available case analysis Study-specific child health questionnaire: Ingestion of poison Follow-up: mean 52 weeks	Study population		RR 0.11 (0.01 to 2.08)	138 (1 study)	⊕⊖⊖⊖ very low2,3,4	
	57 per 1000	6 per 1000 (1 to 119)				
	Moderate					
	57 per 1000	6 per 1000 (1 to 119)				
	Study population					

Child protective service reports (all types) Post-treatment (substantiated reports during trial measured at endpoint or first measurement) – ITT analysis Child protective service reports: Substantiated reports of all types Follow-up: mean 104 weeks	330 per 1000	313 per 1000 (231 to 422)	RR 0.95 (0.7 to 1.28)	364 (1 study)	⊕⊖⊖⊖ very low1,2,3
	Moderate				
	330 per 1000	314 per 1000 (231 to 422)			
Child protective service reports (all types) Post-treatment (substantiated reports during trial measured at endpoint or first measurement) – available case analysis Child protective service reports: Substantiated reports of all types Follow-up: mean 104 weeks	Study population		RR 0.94 (0.57 to 1.56)	297 (1 study)	⊕⊖⊖⊖ very low1,2,3
	173 per 1000	163 per 1000 (99 to 270)			
	Moderate				
	173 per 1000	163 per 1000 (99 to 270)			
Child protective service reports (neglect) Post-treatment (substantiated reports during trial measured at endpoint or first measurement) – ITT analysis Child protective service reports: Substantiated reports of neglect Follow-up: mean 104 weeks	Study population		RR 0.94 (0.68 to 1.3)	364 (1 study)	⊕⊖⊖⊖ very low1,2,3
	297 per 1000	279 per 1000 (202 to 386)			
	Moderate				
	297 per 1000	279 per 1000 (202 to 386)			
Child protective service reports (neglect) Post-treatment (substantiated reports during trial measured at endpoint or first measurement) – available case analysis Child protective service reports: Substantiated reports of neglect Follow-up: mean 104 weeks	Study population		RR 0.92 (0.51 to 1.66)	297 (1 study)	⊕⊖⊖⊖ very low1,2,3
	133 per 1000	123 per 1000 (68 to 221)			
	Moderate				
	133 per 1000	122 per 1000 (68 to 221)			
Maternal use of punishment Post-treatment (corporate/verbal	Study population		RR 0.96 (0.86 to 1.08)	364 (1 study)	⊕⊕⊖⊖ low1,2
	789 per 1000	758 per 1000 (679 to 852)			
	Moderate				

punishment used anytime in past week measured at endpoint or first measurement) – ITT analysis Straus's parent-child Conflict Tactics Scale (CTS-PC): Corporate/verbal punishment Follow-up: mean 104 weeks	789 per 1000	757 per 1000 (679 to 852)				
Maternal use of punishment Post-treatment (corporate/verbal punishment used anytime in past week measured at endpoint or first measurement) – available case analysis Straus's parent-child Conflict Tactics Scale (CTS-PC): Corpoarte/verbal punishment Follow-up: mean 104 weeks	Study population 683 per 1000	656 per 1000 (553 to 785)	RR 0.96 (0.81 to 1.15)	249 (1 study)	⨁⨁⊖⊖ low1,2	
	Moderate 683 per 1000	656 per 1000 (553 to 785)				
Potential for child abuse Post-treatment (mean score at endpoint or first measurement) – available case analysis Child Abuse Potential Inventory Follow-up: mean 78 weeks		The mean potential for child abuse post-treatment (mean score at endpoint or first measurement) – available case analysis in the intervention groups was 0.36 standard deviations lower (0.71 lower to 0 higher)		124 (1 study)	⨁⊖⊖⊖ very low4,5	SMD -0.36 (-0.71 to 0)

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 High risk of selection bias due to unclear allocation concealment and statistically significant baseline differences in poor psychological resources (37% intervention group versus 50% control) and in prenatal enrolment (41% intervention group and 53% control)

2 Total number of events is less than 300 (a threshold rule-of-thumb)
3 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
4 High risk of selection bias due to statistically significant baseline group differences in: parity (54% of intervention group primiparous versus 33% of control); identification as indigenous Australian (9% of intervention versus 2% of control); mental illness of partner (3% of intervention versus 14% of control); history of postnatal depression (11% of intervention versus 28% of control); physical domestic abuse (2% of intervention versus 10% of control); potential for child abuse (mean CAPI score in intervention was 123 versus 159 in control, and elevated CAPI score for 12% of intervention group versus 30% of control group)
5 Total population size is less than 400 (a threshold rule-of-thumb)

7.5.25 Clinical evidence for effects on optimal infant care (by intervention)

Summary of findings can be found in the tables presented in this section. The full GRADE evidence profiles and associated forest plots can be found in Appendix 22 and Appendix 19, respectively.

Optimal infant care: structured psychological interventions (CBT or IPT) versus treatment as usual or enhanced treatment as usual

A single study (N=705-9.3) found no evidence for clinically significant effects (although effects were statistically significant) of CBT relative to enhanced treatment as usual (home visits) on complete immunisation (p=0.04-0.0001) (Table 259).

Table 259: Summary of findings table for effects of structured psychological interventions (CBT or IPT) compared with treatment as usual or enhanced treatment as usual on optimal care of the infant

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Optimal infant care: Structured psychological interventions (CBT or IPT) versus TAU/Enhanced TAU				
Immunisation Post-treatment (complete immunisation at endpoint or first measurement) – ITT analysis Optimal infant care: Complete	Study population		RR 1.1 (1.01 to 1.19)	903 (1 study)	⊕⊕⊕⊕ high	
	668 per 1000	735 per 1000 (675 to 795)				
	Moderate					
	668 per 1000	735 per 1000 (675 to 795)				

immunisation Follow-up: mean 52 weeks					
Immunisation Post-treatment (complete immunisation at endpoint or first measurement) – available case analysis Optimal infant care: Complete immunisation Follow-up: mean 52 weeks	Study population 852 per 1000	946 per 1000 (895 to 989)	RR 1.11 (1.05 to 1.16)	705 (1 study)	⊕⊕⊕⊕ high
	Moderate 852 per 1000	946 per 1000 (895 to 988)			

*The basis for the assumed risk (for example, the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

7.5.26 Health economics evidence

Systematic literature review

The systematic literature search identified three eligible UK studies (Hewitt et al., 2009; Paulden et al., 2009; Morrell et al., 2009a; Stevenson., 2010a [HTA]; Stevenson et al., 2010b) and one Canadian study (Dukhovny et al., 2013) that assessed the cost effectiveness of psychosocial interventions in postnatal women with mental health problems. All four identified studies assessed the cost effectiveness of psychosocial interventions for depression in the postnatal period. Details on the methods used for the systematic search of the economic literature are described in Chapter 3. References to included studies and evidence tables for all economic studies included in the guideline systematic literature review are presented in Appendix 21. Completed methodology checklists of the studies are provided in Appendix 20. Economic evidence profiles of studies considered during guideline development (that is, studies that fully or partly met the applicability and quality criteria) are presented in Appendix 22, accompanying the respective GRADE clinical evidence profiles.

Paulden and colleagues (2009) evaluated the cost-utility of structured psychological therapy and listening visits compared with standard care in women with postnatal mild to severe depression managed in primary care. This treatment model was part of a model which was used to assess the cost-utility of screening for depression in

the postnatal period in primary care in the UK. Hewitt and colleagues (2009) reported the same analysis as part of the Health Technology Assessment report. The time horizon of the analysis was 12 months and the perspective of the NHS and PSS was adopted. The effectiveness data were derived from meta-analysis of RCTs. The study estimated intervention costs including clinical psychologist, health visitor, GP and CPN; and also additional costs associated with standard postnatal care for women with depression in the postnatal care. Costs associated with infant care were not included in the estimation of costs, owing to lack of relevant data. The resource use estimates were based on studies that provided efficacy data and where necessary were supplemented with authors' assumptions. The unit costs were obtained from national sources. The measure of outcome for the economic analysis was the QALY.

The expected mean QALYs per woman were 0.7489, 0.7513 and 0.7036 for the structured psychological therapy, listening visits and standard care groups, respectively. The expected incremental cost (relative to standard care) per woman over 12 months was £792 for structured psychological therapy and £947 for listening visits in 2006-2007 prices. The cost per QALY associated with the structured psychological therapy was £17,480 when compared with standard care which is below NICE's lower cost-effectiveness threshold value of £20,000 per QALY; however when using uplifted cost (to 2013/2014 prices) the ICER goes just above £20,000 per QALY (that is, £20,732). The cost per QALY associated with listening visits was £66,275 when compared with structured psychological therapy. Probabilistic analysis indicated that at WTP of £20,000-£30,000 per QALY the probability that structured psychological therapy is cost effective is 0.504-0.549; the probability that listening visits is the most cost-effective intervention is 0.276-0.414 and the probability that standard care is cost effective is 0.220-0.037. Results suggest that structured psychological therapy is the most cost-effective treatment among those assessed, for women with depression in the postnatal period. Even though listening visits resulted in slightly higher number of QALYs, the considerably higher cost of this strategy resulted in a cost per QALY versus structured psychological therapy that was well above the cost-effectiveness threshold of £20,000-£30,000 per QALY considered to represent value for money.

The analysis was judged by the GDG to be directly applicable to this guideline review and the NICE reference case. This was a UK-based study and the outcome measure of the economic analysis was the QALY; however the utility values were derived from the general population with depression treated with antidepressant medication. The relative effect between structured psychological therapy and listening visits was based on indirect comparisons between treatments, using standard care as the baseline common comparator, due to lack of head-to-head comparisons between the two interventions. Some of resource use was informed by expert opinion; costs associated with infant care were excluded due to the lack of relevant data. Nevertheless, given the limited availability of data this was a well

conducted study and was judged by the GDG to have only minor methodological limitations.

Morrell and colleagues (2009a) assessed the cost effectiveness of listening visits, CBT or standard care. The authors also compared the intervention group as a whole (not differentiating between listening visits and CBT) with standard care. The intervention involved health visitor training in systematically identifying depressive symptoms and delivering psychologically informed sessions based on either CBT or listening visits at GP practice. Standard care was defined as care shared between the midwife and a GP, or otherwise consultant-led care based on clinical need. The study population comprised women with EPDS score ≥12 at 6-weeks after childbirth. The mean baseline EPDS of the study sample was 15.2 (SD 3.0) and their mean age was 31 years. This was an economic evaluation undertaken alongside a cluster randomised RCT (MORRELL2009A) that involved 101 general practices (clusters) in 29 primary care trusts in the UK. The efficacy data was derived from RCT (n=418 at 6 months, n=123 at 12 months). The time horizon of the main analysis was 6 months; secondary analysis reported cost effectiveness at 12 months. The perspective of the NHS and PSS was adopted. The study estimated costs associated with health visitor training, health visitor visits, GP contacts, prescriptions, social worker contacts, MBU, paediatric admissions, community mental health contacts, walk-in centre attendances, emergency department attendances and NHS direct contacts. The resource use estimates were based on data collected alongside the RCT (n=248 at 6 months, n=123 at 12 months), expert opinion and authors' assumptions. The unit costs were obtained from national sources and from the RCT (that is, costs pertaining to health visitor training). The measure of outcome for the economic analysis was the QALY.

At 6 months the mean QALYs gained per woman was 0.026 for the intervention group and 0.023 for the standard care group, a difference of 0.003 QALYs (95% CI, -0.004 to 0.010). The mean cost per woman over 6 months was £339 for the intervention group and £374 for the standard care group in 2003/04 prices, a difference of -£35 (95% CI, -£137 to £67). According to the analysis the intervention group provides better outcome at lower cost, and thus is a dominant intervention when compared with the standard care group at 6 months. Furthermore, according to the probabilistic analysis at WTP of £20,000-£30,000 per QALY the probability that the intervention group is cost effective was just above 0.70. Comparing CBT and listening visits with standard care, CBT resulted in QALY gains of 0.004 (0.027 versus 0.023) and listening visits in 0.002 (0.025 versus 0.023). Similarly, CBT resulted in cost savings of £45 (£329 versus £374) and listening visits of £21 (£353 versus £374) when compared with standard care. As a result, CBT was found to be dominant compared with listening visits and standard care, and at WTP of £20,000-£30,000 per QALY the probability that CBT is cost effective was approximately 0.70.

At 12 months the mean number of QALYs gained per woman was 0.117 for the intervention group and 0.107 for the standard care group, a difference of 0.010 QALYs (95% CI, 0.000 to 0.021). The mean cost per woman over 12 months was £763

for the intervention group and £772 for the standard care group, a difference of -£9 (95% CI, -£177 to £159). According to the analysis the intervention group provides better outcome at lower cost, and thus is a dominant intervention when compared with standard care. At WTP of £20,000-£30,000 per QALY the probability that the intervention group is cost effective was estimated to be just over 0.80. There was no difference between CBT and listening visits at 12 months. Overall the results suggest that psychological interventions are cost effective for women with depression in the postnatal period in the UK.

The analysis was judged by the GDG to be directly applicable to this guideline review and the NICE reference case. This was a UK-based study and the outcome measure was the QALY. QALYs were estimated based on SF-36 data, which were converted into utility scores using the SF-6D algorithm and preferences from the UK general population (Brazier et al., 2002). Some of resource use estimates were based on expert opinion and the authors' assumptions; also some of the costs were trial-specific which may limit the generalisability of the findings. Moreover, the attrition rate was quite high. As a result it may have been underpowered to detect differences between CBT and listening visits at 12 months. Overall, this was a well conducted economic analysis and was judged by the GDG to have only minor methodological limitations.

Stevenson and colleagues (2010b) evaluated the cost-utility of CBT-informed psychoeducation compared with standard care in the UK. Stevenson and colleagues (2010a) reported the same analysis as part of Health Technology Assessment report. CBT-informed psychoeducation entailed one session per week for eight weeks, which was of two hour duration and was held in groups of four to six women. Standard care was defined as routine primary care that included visits by midwives and health visitor, GP care, medication, community mental health contacts and social services. This was an economic evaluation based on a small RCT (HONEY2002) (n=45) and modelling. The study population comprised women with EPDS ≥ 12; the mean baseline EPDS of the study sample was 19.5 (SD 4.17). Efficacy data were taken from the RCT. The RCT provided efficacy data at baseline, end of treatment (that is, 8 weeks), and at 6-month follow-up. Based on clinical advice, it was assumed in the base-case analysis that the incremental gain in EPDS of CBT-informed psychoeducation compared with standard care would rise linearly to a peak value at 8 weeks (that is, at the end of intervention), stay constant until 6 months, and then decline linearly to zero by 12 months after randomisation (that is, it was assumed that no effect is retained at 12 months). The incremental gain was assumed to decline to zero at 12 months because symptoms of depression were no longer assumed to be postnatal in origin by that time point. The time horizon of the analysis was 12 months and the perspective of the NHS and PSS was adopted. It was assumed that standard care costs were the same across both groups; consequently the authors estimated only the costs associated with the provision of CBT-informed psychoeducation. The resource use estimates were based on the RCT, other published studies and authors' assumptions. The unit costs were obtained from published studies. The measure of outcome for the economic analysis was the

QALY. In order for QALYs to be estimated a mapping technique was utilised. To do this data was obtained from the PoNDER trial (Morrell et al., 2009a), which collected data on both EPDS and SF-36; the statistical relationship between EPDS and SF-36 and the SF-6D algorithm that converts SF-36 into utility values (Brazier et al., 20024) were subsequently used to transform the observed gains in EPDS recorded in HONEY2002 RCT into utility values that could be utilised in the economic model.

The pooled comparative advantage in EPDS was estimated to be 3.98 points (95% CI, 0.23 to 6.73) in favour of the intervention. Using the mapping technique it was estimated that CBT-informed psychoeducation resulted in a QALY gain of 0.032 (95% CI, 0.025 to 0.041). The incremental cost associated with CBT-informed psychoeducation over 12 months was £1,500 per woman. The cost year of the analysis was 2007/08. The ICER associated with CBT-informed psychoeducation was estimated to be £46,462 per QALY gained (95% CI, £37,008 to £60,728). The sensitivity analysis showed that when the cost of intervention per woman was decreased to £750 (that is, a reduction of 50%), the ICER decreased to £23,231 per QALY; and when the cost of intervention was increased to £2,000 per woman, the ICER increased to £61,948 per QALY. Using the lower estimate of efficacy (that is, EPDS advantage of 3.27 in favour of intervention) the cost per QALY increased to £56,626 and using an upper estimate (that is, EPDS advantage of 4.69 in favour of intervention) it was £39,481. Moreover, assuming a linear decline in advantage of CBT-informed psychoeducation extended to 18 months (instead of the 12 months assumed in the base-case analysis), the resulting ICER became £34,382 per QALY; assuming a QALY gain associated with CBT-informed psychoeducation of 0.02 per woman resulted in a cost per QALY of £28,846. The authors also conducted a scenario analysis where the cost of intervention per woman was decreased to £1,000, the change in EPDS scores was assumed to be 4.3 in favour of CBT-informed psychoeducation, and a linear decline in advantage of group CBT was extended to 18 months. The scenario resulted in a cost per QALY of £19,230 which is just below NICE's lower cost-effectiveness threshold value. Considering the results of the various scenarios explored in sensitivity analysis, the authors concluded that their findings were too uncertain to draw any firm conclusions on the cost effectiveness of CBT-informed psychoeducation in women with depression in the postnatal period.

Nevertheless, the base-case analysis and majority of scenarios explored suggest that CBT-informed psychoeducation is unlikely to be cost-effective intervention in women with depression in the postnatal period at 12 months since the cost per QALY is well above NICE cost-effectiveness threshold of £20,000-£30,000 per QALY considered to represent value for money. Also, the GDG considered that the exclusion of set-up costs and additional running costs such as crèche facilities potentially underestimated the costs associated with the intervention. Nevertheless, the actual cost of CBT-informed psychoeducation based on the resource utilisation reported in RCT was £1,317 and based on the resource use estimates deemed most appropriate by the authors' expert opinion it was £1,246. Moreover, the authors considered only interventions costs, and ignored potential cost-savings resulting from a reduction in depression symptoms.

The analysis was judged by the GDG to be directly applicable to this guideline review and the NICE reference case. This was a UK-based study and outcome measure used was the QALY. QALYs were estimated using mapping technique. Moreover, the estimate of relative treatment effect was obtained from a single small RCT and the authors made a series of assumptions regarding the efficacy of CBT-informed psychoeducation beyond the duration of the RCT. Similarly, the resource use was based on the same small RCT and where necessary it was supplemented with the authors' assumptions. Nevertheless, the authors partially addressed these limitations by conducting extensive sensitivity analyses. Overall, this study was judged by the GDG to have potentially serious methodological limitations.

In a recent study Dukhovny and colleagues (2013) assessed the cost effectiveness of social support (that is, telephone-based peer support service) compared with standard care for women at high-risk for depression in the postnatal period. However, since all of the women in RCT scored >9 on the EPDS and 39% scored >12 the study was classified as treatment study for this guideline review, even though the authors aimed the intervention to be preventative. This was an economic evaluation undertaken alongside an RCT (DENNIS2009) (n=612) conducted in Canada. Social support entailed peer volunteers making a minimum of four telephone contacts initiated 48 to 72 hours after randomisation and continuing through the first 12 weeks after childbirth. Standard care was defined as mother proactively seeking services from public health nurses, physicians, other providers, and various community resources, including drop-in centres. The time horizon of the analysis was 12 weeks and a societal perspective was adopted; however the authors reported costs for different cost categories separately, which enabled estimation of costs from a healthcare perspective. The study estimated public health costs, volunteer opportunity cost, hired housework, hired child care, family/friend and partner time off work, nursing visits, provider visits, mental health visits, and inpatient admissions. The resource use estimates were based on the RCT (n=610) and the unit costs were obtained from local and national sources. The authors used number of cases of depression avoided as an outcome in their economic analysis; however since this study was classified as treatment study for this guideline review the outcome was redefined as number of cases with EPDS score ≤12.

Intervention resulted in a greater proportion of cases with EPDS score ≤12. Percentage of women with EPDS score of ≤12 was 87% and 75% in the intervention and standard care groups, respectively (difference of 11%, $p < 0.05$). The costs in the study were measured in Canadian dollars in 2011 prices. From a healthcare payer perspective the mean cost per mother–infant dyad over 12 weeks was $1,694 for the intervention and $1,080 for standard care, difference of $614. From a societal perspective the mean cost per mother–infant dyad over 12 weeks was $4,497 for the intervention and $3,380 for standard care, difference of $1,117 ($p<0.05$). The cost per additional woman with EPDS score ≤12 was $10,009 and $5,582 from a societal perspective (plus informal care) and a healthcare payer perspective, respectively. Sensitivity analysis was conducted only on the results from a societal perspective. As

the number of healthcare visits was varied between 50% and 400% of the number used in the base-case analysis, the ICER ranged from $9,671 to $9,110 per additional case with EPDS score of ≤12. The ICER was most sensitive to the cost of running the programme, volunteer time, family/friend and partner work absence. Moreover, probabilistic analysis showed that at WTP of $20,196 per case with EPDS score of ≤12 the probability of the intervention being cost effective was 0.95. Results suggest that intervention provides better outcomes but at an additional cost.

The analysis was judged by the GDG to be partially applicable to this guideline review and the NICE reference case. The study was conducted in Canada where the healthcare system is sufficiently similar to the UK NHS. The authors did not attempt to estimate QALYs which made it difficult to interpret the cost effectiveness results and to compare the findings with other studies. Also, a mixture of local and national unit costs were utilised which may limit the generalisabiltiy of the findings to other settings. Moreover, the effectiveness was based on one RCT and the time horizon was only 12 weeks which may not be sufficient to reflect all important differences in costs and outcomes. Also, the sensitivity analysis was conducted only on the results derived using a societal perspective. As a result, the study was judged by the GDG to have potentially serious methodological limitations.

Overall conclusions from existing economic evidence

The existing economic evidence on psychological and psychosocial interventions for the treatment of mental health problems in women who are pregnant or in the postnatal period is very sparse and limited to depression in the postnatal period. The systematic literature search identified three UK-based economic evaluations that were all judged by the GDG to be directly applicable to the NICE decision-making context. Two of the studies included in the review were characterised by minor methodological limitations and one by potentially serious limitations. In one of the studies the structured psychological therapy was found to be cost-effective option when compared with standard care, as it resulted in an ICER of £17,480 per QALY; however when using uplifted cost (to 2013/2014 prices) the ICER goes just above £20,000 per QALY. In another study psychological therapy resulted in better outcomes at lower cost, and thus was found to be dominant when compared with standard care. The third study indicated that CBT-informed psychoeducation was not cost effective compared with standard care. The results of the Canadian study were inconclusive, as they do not use QALYs and it is difficult to judge whether the reported extra benefits associated with the intervention are worth the extra costs associated with its provision.

Economic modelling

Introduction – objective of economic modelling

The provision of psychological and psychosocial interventions aimed at treating depression during postnatal period in women with subthreshold/mild to moderate depression was identified by the GDG as an area with potentially significant resource implications. The existing economic evidence was not sufficient to support

decision making by the GDG, consequently a decision-analytic model was developed to assess the cost effectiveness of different types of psychological and psychosocial interventions added to standard postnatal care, relative to standard postnatal care alone, for the treatment of depression in the postnatal period.

The study population

The study population consisted of women with subthreshold/mild to moderate depression in the postnatal period.

Economic modelling methods

Interventions assessed

The economic model considered interventions that were found to be effective in the meta-analysis conducted for this guideline. Two different types of treatments were considered:

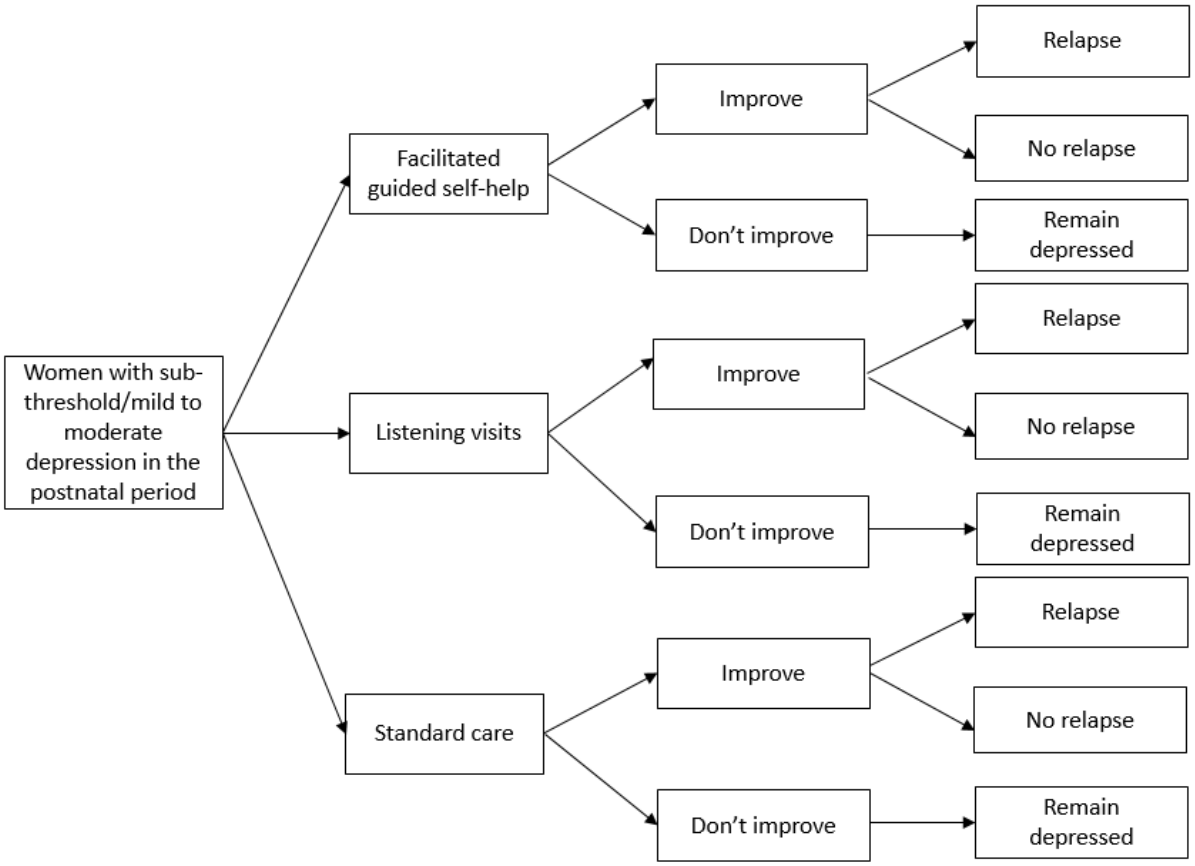
- facilitated self-help added to standard postnatal care
- listening visits added to standard postnatal care

In addition, standard postnatal care alone was considered as an alternative option, in order for the active treatments to be assessed.

Model structure

The economic model was developed in the form of a decision tree using Microsoft Office Excel 2013. According to the model structure, hypothetical cohorts of 1,000 women with subthreshold/mild to moderate depression in the postnatal period received one of the treatments assessed. At the end of treatment (that is, 7 weeks), women either improved or did not improve. Women were followed for 1 year since initiation of treatment. Over this period, women who improved, either remained in this state or relapsed. Responders to treatment in each trial that provided efficacy data for the model were calculated on an intention-to-treat basis (that is, response rates were estimated for those who were randomised in each arm and not only for those who completed treatment); consequently discontinuation has not been considered separately in the model. A schematic diagram of the decision-analytic model is presented in Figure 11.

Figure 11: Schematic diagram of the structure of the economic model



Costs and health benefit measures included in the analysis

The analysis adopted the perspective of the NHS and PSS. Costs consisted of treatment costs (facilitated self-help or listening visits), and health and social care costs for mother–infant dyad. Standard postnatal care costs were omitted from the analysis, because they were common to all therapeutic options assessed. Other costs to women and family, such as personal expenses and productivity losses were also excluded as they were beyond the scope of the analysis. Intangible costs (negative impact of the woman's depression on infant's cognitive and emotional development as well as distress to the family) were also not estimated, but they should be taken into account when interpreting the results.

Two different measures of health benefits were used in the economic analysis:

1. Number of women who improved and did not relapse at the end of 1-year follow-up
2. Number of quality adjusted life years (QALYs) gained at the end of 1-year follow-up.

Total costs and health benefits associated with each treatment were estimated and combined in order to assess the relative cost effectiveness of the treatment options evaluated.

Effectiveness data and other input parameters of the economic model

Effectiveness data used in the economic model were derived from the guideline meta-analyses. All studies providing dichotomous efficacy data on facilitated self-help and listening visits in the study population were considered in the economic analysis. The types of treatments examined in each of the studies considered are presented in Table 260.

Table 260: Types of treatments of depression in the postnatal period examined in the clinical studies considered in the economic analysis

Study	*Treatments assessed (in addition to standard care)*
MILGROM2011A	Facilitated self-help that included towards parenthood intervention and community networking delivered over 8 weeks; self-help book
OMAHEN2013A	Facilitated self-help (computer-delivered CBT); 11 (internet sessions) and 1-2 (median support sessions) delivered over 15 weeks
OMAHEN2013C	Facilitated self-help delivered (computer-delivered CBT) over 8 computer sessions; a mean of 8 telephone support sessions
MORRELL2009A/2009B/2011	Eight individual weekly listening visits delivered by health visitors trained in Person Centred Approach
WIGGINS2005	Ten individual listening visits delivered by very experienced health visitors

Since there were no direct comparisons between the treatments under assessment, it was decided to perform an indirect comparison between them. In order to do this, relative risks of non-improvement (efficacy) of each of the two treatments versus standard care were used, with standard care serving as the baseline common comparator. The absolute rate of non-improvement associated with standard care were based on the whole dataset of studies evaluating treatments for depression in the postnatal period, included in the guideline systematic review, that had a 'standard care' arm (that is, all studies reported in Table 260).

The absolute risks of non-improvement of each treatment were estimated by multiplying the respective relative risks for each treatment, derived from meta-analysis, by the absolute risk of non-improvement as calculated for standard care, using the formula:

$$NIAR_{int(i)} = NIRR_{int(i)} \times NIAR_{st\ care}$$

where:

$NIAR_{int(i)}$ = absolute risk of non-improvement of each treatment

$NIRR_{int(i)}$ = relative risk of non-improvement of each treatment versus standard care

$NIAR_{st\ care}$ = absolute risk of non-improvement of standard care

It is acknowledged that the indirect comparison between treatments may have introduced some degree of bias in the analysis, as there were differences between the studies in terms of severity of depression in study samples, diagnostic measures used, content of treatments and comparators, and some other aspects of protocol

design. Nevertheless, due to the limited availability of data, the indirect comparison was considered necessary in order to populate the economic model.

Estimation of relapse risk

The risk of relapse over 12 months was assumed to be common to women improving following treatment as well as to women having improved under standard care. No studies reporting relapse rates for the study population were identified. As a result it was assumed that a mean of 50% of women would relapse over 12 months. Relapse rates were utilised in the model for the estimation of benefits in the form of QALYs and also in the estimation of additional costs due to relapse.

Utility data and estimation of QALYs

Similarly to the economic model described in Chapter 5 (section 5.3.6), utility values for this economic analysis were taken from the study by Sapin and colleagues (2004). Utility scores for 'subthreshold/mild to moderate' depression in the model were approximated using utility scores reported in Sapin and colleagues (2004) for 'slightly/moderately ill'. Based on the GDG expert opinion 'no depression' health state in the model was approximated using utility scores for 'first signs' depression reported in the study; the value of which was also very similar to utility scores reported for 'responder remitters'.

The use of these data in the cost-utility analysis performed for this guideline is characterised by a number of limitations:

- Data express the HRQoL of the general population of service users with depression and are not specific to women with depression in the postnatal period. However, this period is associated with wide physical and emotional events in women's lives, which are likely to further affect their HRQoL.
- Data refer to utility weights of service users under antidepressant medication, and therefore may incorporate aspects of treatment such as the presence of side effects that are not relevant to the treatments examined in this analysis.
- Data refer to women's HRQoL, and they do not take into account that of the babies, which is subsequently affected by their mother's psychological condition. Although, it would be very difficult to actually measure the babies HRQoL and express it in utility weights, this parameter should be considered in the interpretation of the results.

In the model women who improved were assumed to experience a linear improvement in their HRQoL (expressed in QALYs) from initiation to the end of treatment. Women who relapsed within the first year were assumed to experience a linear deterioration in their HRQoL from the time of relapse until the model endpoint. Women who have not improved where assumed to remain in their original health state (that is, depressed health state) until the model endpoint.

All effectiveness rates and other input parameters included in the economic model are provided in Table 261.

Cost data

Since no patient-level data in terms of resource use were available, the economic analysis was based on deterministic costing of the treatment options. Relevant healthcare resource use was estimated and subsequently combined with UK unit prices to provide costs associated with each treatment strategy assessed. Estimated resource use associated with the two treatments evaluated (facilitated self-help and listening visits) was based on definitions of the treatments in the studies that provided the efficacy data. Further healthcare resource use required was based on the GDG expert opinion, owing to lack of research-based evidence.

Petrou and colleagues (2002) estimated the economic costs of depression in the postnatal period in a geographically defined cohort of women at high risk of developing the condition. Health and social care costs were estimated based on 206 women recruited from antenatal clinics and their babies. The study estimated costs associated with community care, day care services, hospital outpatient attendances, hospital inpatient admissions, and paediatric and child care services. The reported health and social care costs for women with depression in the postnatal period were utilised in the model to estimate health and social care costs associated with women who haven't improved or those who have relapsed. Similarly, women who have improved were assigned health and social care costs associated with women with no depression in the postnatal period.

Unit prices were taken from national sources (Curtis, 2013). All costs utilised in the analysis reflect 2013-2014 prices. Discounting of costs was not applied, as the time horizon of the analysis was 1 year and 7 weeks. Table 118 shows the estimated resource use and total costs associated with each treatment option.

Table 261: Effectiveness data and other input parameters included in the model

Input parameter	Deterministic value	Probabilistic distribution	Source of data — comments
Clinical input parameters			
Relative risk of non-improvement Facilitated self-help Listening visits	 0.73 0.96	**Log-normal distribution** 95% CI, 0.53 to 0.99 95% CI, 0.84 to 1.09	Guideline meta-analysis
Absolute risk of non-improvement Standard care	 0.61	**Beta distribution** α=793, β=508	Guideline meta-analysis
Relapse risk at 12 month follow-up	0.50	**Beta distribution** α=50, β=50	GDG expert opinion; distribution parameteres based on assumption
Utility scores No depression Subthreshold/mild to moderate depression in the postnatal period	 0.86 0.74	**Beta distribution** α=86, β=14 α=74, β=26	Sapin and colleagues (2004); utility scores for the general depression population treated with antidepressant medication; utility score for slightly/moderately ill reported by Sapin and clolleagues (2004) was used as a proxy for subthreshold/mild to moderate depression in the postnatal period; distribution parameters based on assumption
Cost data (2013/2014 prices)			
Intervention cost Facilitated self-help	£224.92	**Gamma distribution** α=11, β=20	Based on seven telephone-based support sessions (25 minutes per session) provided by psychological wellbeing practitioner (Band 5) trained in perinatal issues; plus facilitated self-help manual costing £9.09 (*Overcoming Depression: A 'Books on Prescription' Title: A Self-help Guide Using Cognitive Behavioral Techniques* by Paul Gilbert; Amazon.co.uk). Unit cost of psychological wellbeing practitioner unavailable; unit cost approximated using unit cost of mental health nurse (Band 5) £74 per hour (Curtis, 2013). To estimate probabilistic distribution standard error assumed to be 30% of its mean estimate because of a lack of data.

Intervention cost Listening visits	£497.00	**Gamma distribution** α=11, β=45	Based on seven, weekly health visitor home visits × 60 minutes each session (studies in guideline meta-analysis and GDG expert opinon). Unit cost of health visitor £71 per hour of home visiting (Curtis, 2013). To estimate probabilistic distribution standard error assumed to be 30% of its mean estimate because of a lack of data.
Weekly health and social care costs Women with depression in the postnatal period Women with no depression in the postnatal period	£50.66 £42.52	**Gamma distribution** α=11, β=5 α=11, β=4	Petrou and colleagues (2002); costs reported were uplifted to 2013/14 UK pounds using UK HCHS inflation index.

Handling uncertainty

In order to take into account the uncertainty characterising the model input parameters, a probabilistic analysis was undertaken, in which input parameters were assigned probability distributions, rather than being expressed as point estimates (Briggs et al., 2006). Subsequently, 1000 iterations were performed, each drawing random values out of the distributions fitted onto the model input parameters. Mean costs and QALYs for each intervention were then calculated by averaging across 1000 iterations.

The relative risk of non-improvement associated with facilitated self-help and listening visits were given a log-normal distribution. The absolute risk of non-improvement were given a beta distribution. Beta distributions were also assigned to utility values and relapse rate. Costs were assigned a gamma distribution. The estimation of distribution ranges was based on available data in the published sources of evidence, and further assumptions where relevant data were not available. Table 261 provides details on the types of distributions assigned to each input parameter and the methods employed to define their range.

One-way sensitivity analyses (run with the point estimates rather than the distributions of the input parameters) explored the impact of the uncertainty characterising the model input parameters on the model's results:

- changes in relative risk estimates
- changes in the absolute risk of non-improvement associated with standard care
- changes in utility weights
- changes in treatment costs

Moreover, threshold sensitivity analyses were also conducted to explore the magnitude of change in base-case values of input parameters required for the conclusions from cost-utility analysis to be reversed.

Data analysis and presentation of the results

Results of the economic analysis are presented as follows:

For each intervention mean total costs, number of women improving and not relapsing at the end of model, and QALYs are presented, averaged across 1000 iterations of the model. An incremental analysis is provided, where all options have been ranked from the most to the least effective (in terms of QALYs gained). Options that are dominated by absolute dominance (that is, they are less effective and more costly than one or more other options) are excluded from further analysis. Subsequently, Incremental Cost Effectiveness Ratios (ICERs) are calculated for all pairs of consecutive options remaining in analysis.

ICERs are calculated by the following formula:

$$ICER = \Delta C / \Delta E$$

where ΔC is the difference in total costs between two interventions and ΔE the difference in their effectiveness (QALYs). ICERs express the extra cost per extra unit of benefit (that is, QALY in this analysis) associated with one treatment option relative to its comparator. The treatment option with the highest ICER below the NICE lower cost-effectiveness threshold of £20,000 per QALY (NICE, 2008) is the most cost-effective option.

Moreover, for the most cost-effective intervention, the probability that this is the most cost-effective option is also provided, calculated as the proportion of iterations (out of the 1000 iterations run) in which the intervention was the most cost effective among all interventions considered in the analysis.

Validation of the economic model

The economic model (including the conceptual model and the excel spreadsheet) was developed by the health economist working on this guideline and checked by a second modeller not working on the guideline. The model was tested for logical consistency by setting input parameters to null and extreme values and examining whether results changed in the expected direction. The results were discussed with the GDG for their plausibility.

Economic modelling results

Results of the probabilistic analysis are presented in Table 262. Facilitated self-help dominated listening visits as it resulted in more women who have improved and not relapsed at the end of model, in greater gains in QALYs and at the same time it was also less costly. Facilitated self-help compared with standard care was overall more effective and more costly. The ICER of facilitated self-help was £2,269 per additional woman improving and not relapsing at the end of the model, or £13,324 per QALY gained, which is well below NICE's cost-effectiveness threshold of £20,000-£30,000 per QALY gained, indicating that facilitated self-help is likely a cost-effective option compared with standard care. The cost-effectiveness plane showing the incremental costs and QALYs of facilitated self-help versus standard care, facilitated self-help versus listening visits and listening visits versus standard care resulting from 1000 iterations of the model is shown in Figure 12. The probability of facilitated self-help being cost effective at the NICE cost-effectiveness threshold of £20,000-£30,000 per QALY is 0.59 to 0.72. In Figure 13 cost-effectiveness acceptability curve is presented showing the probability of facilitated self-help being cost effective at various threshold values.

Table 262: Results of the probabilistic analysis referring to a hypothetical cohort of 1,000 women with subthreshold/mild to moderate depression in the postnatal period

Treatment option	QALYs gained	Number of women improving and not relapsing at the end of model	Costs (£)	Incremental QALYs (versus standard care)	Incremental costs (£) (versus standard care)	Cost effectiveness
Facilitated self-help	789	277	£2,358,648	14	£181,117	ICER versus standard care: £2,269 per additional woman improving and not relapsing; £13,324 per QALY gained
Listening visits	764	213	£2,663,386	-	-	Dominated by facilitated self-help
Standard care	775	197	£2,177,530	-	-	

Figure 12: Cost-effectiveness plane showing incremental costs and QALYs of facilitated self-help versus standard care, facilitated self-help versus listening visits, and listening visits versus standard care (per woman). Results based on 1000 iterations

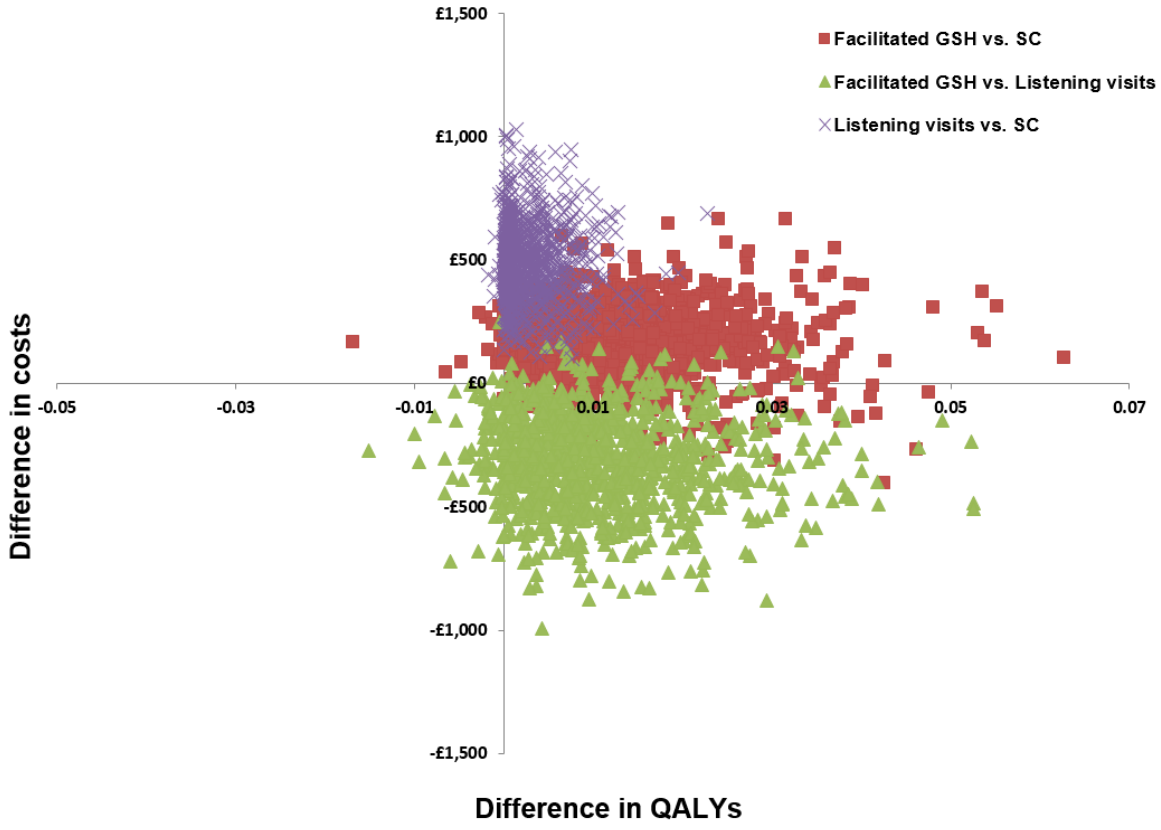
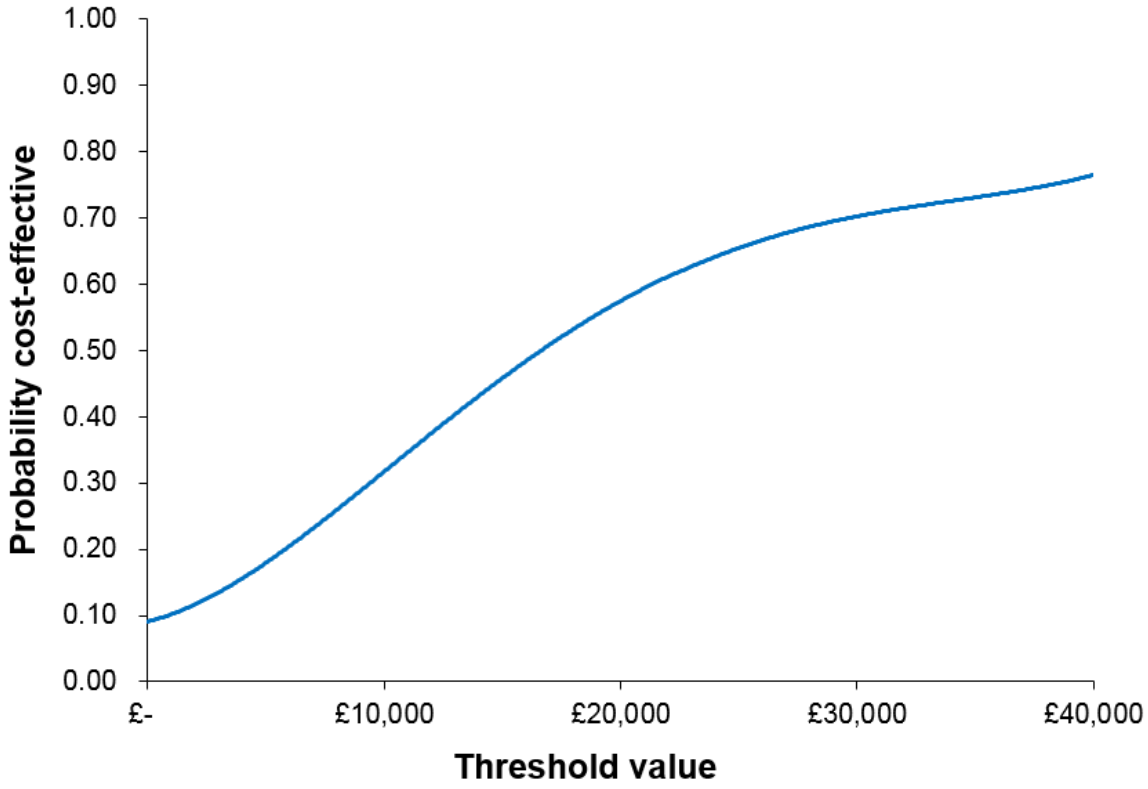


Figure 13: Cost-effectiveness acceptability curve showing the probability of facilitated self-help being cost effective at various threshold values



One-way sensitivity analyses showed that increasing the relative risk of non-improvement associated with facilitated self-help by approximately 20% (from the base-case value of 0.73 to 0.87) would increase the cost per QALY associated with facilitated self-help (relative to standard care) to £29,797 per QALY which is just below NICE's upper cost-effectiveness threshold of £30,000 per QALY. Moreover, only if the relative risk of non-improvement associated with listening visits was reduced to 0.50 (from the base-case value of 0.96), listening visits would be the preferred treatment option with cost per QALY of £19,353 (when compared with facilitated self-help). As the absolute risk of no improvement (that is, 0.61) associated with standard care is varied the conclusions do not change. Only, if it is as low as 0.25 the standard care would become the preferred option; however this would imply the spontaneous recovery rate (rate of improvement) associated with standard care of 0.75 which is unrealistic in clinical practice. Also, if the utility value associated with subthreshold/mild to moderate depression was increased to 0.81 the ICER of facilitated self-help versus standard care would be above NICE's upper cost-effectiveness threshold, and standard care would be the preferred option (that is, an ICER of £30,420 per QALY). In a scenario where treatment costs were varied by 50% either way of their base-case estimates the conclusions did not change. Overall sensitivity analysis indicates that the conclusions of this analysis are very robust to changes in the model's inputs, and only large changes in the base-case values would be required for the model's conclusions to change.

Discussion – limitations of the analysis

Based on the results of the economic analysis, it can be concluded that facilitated self-help is likely to be a cost-effective treatment option for women with subthreshold/mild to moderate depression in the postnatal period. Facilitated self-help was found to be dominant when compared with listening visits, and resulted in an ICER of £13,324 per QALY gained when compared with standard care. The probability of facilitated self-help being cost effective at the NICE cost-effectiveness threshold of £20,000-£30,000 per QALY was 0.59 to 0.72.

Results were driven by the superior efficacy (expressed by the relative risk of non-improvement) of facilitated self-help and the relatively low intervention costs. It should be noted that clinical benefits from treatment are expected to be higher than those estimated in the analysis, since improvement in women's psychological condition has a significant positive impact on babies' cognitive and emotional development, as well as on the well-being of their wider family.

The economic analysis was undertaken using the most accurate effectiveness and cost data available. However, evidence on clinical effectiveness was based on indirect comparisons between treatments, derived from a very limited number of studies. Cost estimates were based on the description of relevant healthcare resource use as provided in the clinical studies, further supported by the GDG opinion.

Utility weights used in the model referred to HRQoL of the general population of service users with depression and not women with depression in the postnatal period. The quality of life of babies and of the wider family associated with the mother's development of depression in the postnatal period was not addressed in the analysis, as relevant data weren't available.

It is recognised that, overall, results of the analysis are subject to some uncertainty regarding some input parameters and potential bias; nevertheless as indicated by the extensive sensitivity analysis the conclusions are robust to changes in model's inputs.

Further research is needed on the efficacy and acceptability of psychological and psychosocial treatments for the management of women with depression in the postnatal period, on the HRQoL of women with this condition and their babies, and on the long-term costs of health and social care of those babies, in order to determine more accurately the relative cost effectiveness of psychological treatments and assist decision making.

Overall conclusions from economic evidence

The existing economic evidence on psychological and psychosocial interventions for the treatment of mental health problems in pregnancy or the postnatal period is very sparse and limited to depression. Even though the search has identified three UK-based economic evaluations that were all judged by the GDG to be directly applicable to the NICE decision-making context, the studies have not looked at the

interventions that were found to be clinically effective in the meta-analysis conducted for this guideline review. In the economic analysis conducted for this guideline, low cost interventions such as facilitated self-help appear to be more cost-effective options than listening visits or standard care. However, the analysis has not overcome many of the limitations characterising previous studies conducted in the area. For example clinical effectiveness was based on indirect comparisons between treatments, derived from a very limited number of studies, some of the resource use estimates were based on the GDG expert opinion and utility values were for the general population with depression. The aforementioned limitations should be considered when making recommendations.

7.6 LINKING EVIDENCE TO RECOMMENDATIONS

In reviewing the evidence for psychosocial interventions aimed at mental health problems in pregnancy and/or the postnatal period the GDG were guided by the principle that much of the treatment of mental health problems in pregnancy and the postnatal period is not different from that at other times of a woman's life, and so should be guided by relevant NICE guidelines for the specific mental health problem. However, new recommendations were developed where there was new evidence specifically for this guideline:

- for an intervention that was specific to pregnancy or the postnatal period;
- that an existing recommendation needed to be clarified or modified as a result of concerns about the health of the fetus or infant;
- that changes are necessary to the context in which interventions are delivered;
- that specific variations are necessitated by changes in a woman's mental or physical health linked to pregnancy and the postnatal period.

In line with these principles, the GDG identified the change to the risk-benefit ratio when considering pharmacological and psychosocial treatments as an instance which necessitated modification to existing guidance for women who are planning a pregnancy, are pregnant, or are breastfeeding. Moreover, the GDG felt that it was a key priority that treatment decisions and discussions be informed by a consideration and trade-off of risks associated with changing or stopping medication during pregnancy (see Chapter 8), the higher threshold for pharmacological interventions due to potential teratogenic harms (see Chapter 8), and the greater prioritisation of prompt and effective psychological interventions. The GDG were particularly mindful that in cases where the optimal treatment is combined psychosocial and pharmacological treatment, but the woman declines or stops taking medication, it is important that she is adequately supported and has the opportunity to discuss the risk associated with stopping psychotropic medication and is offered, or can continue with, a psychological intervention.

These principles also guided the GDG in the decision to restrict the inclusion criteria for study design to RCTs, and exclude observational studies, for the review of treatment efficacy. It was considered appropriate to restrict review to the highest

level of the evidence hierarchy so as to enable consistent linking with other NICE guidance based on wider populations.

Crucial to the effective delivery of any psychosocial intervention is the competence of the staff who are delivering it, and non-adherence with treatment models is associated with a significant attenuation in treatment effects. The GDG reviewed the recommendation from the guideline on depression in adults (NICE, 2009a) and agreed with the need for effective supervision and process-and-outcome monitoring and accordingly adapted the recommendation for women with mental health problems in pregnancy or the postnatal period. The GDG also stressed the importance of prompt delivery and highlighted this as another instance where existing recommendations needed to be modified as more urgent intervention may be required in pregnancy or the postnatal period (than would usually be the case) because of the potential effect of the untreated mental health problem on the fetus/baby and on the woman's physical health and care, and her ability to function and care for her family. The GDG reviewed the previous 2007 recommendation which specified that psychological treatment should be initiated within 1-3 months post-assessment and expressed concerns that women may be placed on waiting lists for assessment so that waiting times for treatment may be considerably longer than the 1-3 month time period outlined. In order to remove this potential ambiguity and ensure prompt delivery, the GDG recommended time scales for assessment (assess for treatment within 2 weeks of referral) and treatment initiation (provide psychological interventions normally within 1 month of initial assessment).

There was very low to high quality evidence from up to three studies for moderate clinical benefits of facilitated self-help on depression symptomatology (scoring above threshold on a depression rating scale) and mean depression symptoms for women with subthreshold to moderate symptoms of depression in pregnancy or the postnatal period. The economic analysis conducted for this guideline also found facilitated self-help to be dominant when compared with listening visits, and result in an ICER of £13,324 per QALY gained when compared with standard care. The probability of facilitated self-help being cost effective at the NICE cost-effectiveness threshold of £20,000-£30,000 per QALY was 0.59 to 0.72. Results were driven by the superior efficacy of facilitated self-help and the relatively low intervention costs. The GDG considered this evidence together with what is known about the clinical and cost effectiveness of facilitated self-help for the treatment of depression in non-pregnant women, and recommended that facilitated self-help should be considered for women with persistent subthreshold depressive symptoms, or mild to moderate depression, and delivered as described in recommendation 1.4.2.2 of the guideline on depression in adults (NICE, 2009a), including the provision of written materials, supported by a trained practitioner (face-to-face or by telephone) and typically consisting of six to eight sessions over nine to twelve weeks.

There was very low to high quality evidence from up to ten studies for large to moderate benefits of structured psychological interventions (CBT or IPT) on depression diagnosis, depression symptomatology and depression mean symptoms,

and some low quality evidence for maintained moderate to large effects at short-term and intermediate follow-up periods. There was also low quality, single study evidence for a large effect of structured psychological interventions on mean anxiety symptoms for women with a diagnosis or symptoms of depression. The economic evidence review also suggested that structured psychological interventions may be cost effective. In the UK studies reviewed structured psychological therapy resulted in a cost per QALY that was within NICE's cost-effectiveness threshold values of £20,000-£30,000 per QALY (when compared with standard care) or was the dominant intervention. Moreover at WTP of £20,000-£30,000 per QALY structured psychological therapy had a greater than 50% probability of being a cost-effective strategy. The GDG considered this evidence together with the much larger evidence base for the clinical and cost effectiveness of structured psychological interventions for the treatment of depression in non-pregnant populations, and took the view that women with moderate to severe depression in pregnancy or the postnatal period should be offered a range of options in line with existing NICE guidance. In adapting existing NICE guidance the GDG took into account the higher threshold for pharmacological intervention for pregnant or breastfeeding women. The range of treatment options include structured psychological interventions alone, pharmacological interventions alone (providing the woman understands the risks and expresses a preference), or combined structured psychological (CBT or IPT) interventions and psychotropic medication in the case of a limited response to either psychological or pharmacological interventions alone. For the evidence for pharmacological interventions and decisions regarding recommendations specifically about drug treatment see Chapter 8.

There was limited evidence for the effectiveness of a pre-delivery psychoeducational discussion on fear of childbirth (symptoms of tokophobia). There were no clinically or statistically significant effects on mode of delivery. However, there was single study evidence for small and statistically significant benefits of pre-delivery discussions on continuous measures of feeling safe during childbirth, the experience of fear during childbirth, and maternal attitude to motherhood. The economic evidence review did not find any studies assessing the cost-effectiveness of pre-delivery interventions for tokophobia. Although the evidence for large and appreciable benefits was not found, the GDG agreed by consensus judgement, that it is important for women with tokophobia to have the opportunity to discuss these fears during the pre-delivery period and they should have access to a healthcare professional with expertise in providing perinatal mental health support. Moreover, the GDG judged that the cost of such interventions would be small relative to the reduction in women's potential for developing mental health problems and other health vulnerabilities which may be costly to other parts of the NHS. Moreover, this recommendation is in line with NICE guidance on caesarean section (NICE, 2011e).

There was no evidence for statistically or clinically significant benefits (or harms) associated with post-traumatic birth counselling on PTSD outcomes for women who had a diagnosis of PTSD. Based on this inconclusive evidence base there were no grounds for recommending postnatal-specific intervention and the GDG

recommended that women with PTSD which has resulted from a traumatic birth, miscarriage, stillbirth or neonatal death should be treated in line with the guideline on PTSD (NICE, 2005a). The GDG reviewed the recommendation from the previous 2007 guideline and judged that the term 'single-session formal debriefing' may be misinterpreted as it is used to refer to post-delivery discussions (without an explicit focus on 're-living' the traumatic experience) in an obstetric context, therefore the decision was taken to modify the previous recommendation and replace the term 'formal debriefing' with 'high-intensity psychological interventions with an explicit focus on 're-living' the trauma'. The GDG also wished to highlight that this guidance applied to women who had experienced a birth as psychologically traumatic (even when the delivery is obstetrically straight forward), and that this guidance also applied for women experiencing trauma symptoms in a pregnancy subsequent to a traumatic birth, miscarriage, stillbirth or neonatal death.

There was no other evidence for the treatment of anxiety disorders in pregnancy or the postnatal period. NICE guidelines for specific anxiety disorders did not specifically include recommendations for people with persistent subthreshold symptoms of anxiety. However, the GDG considered the evidence base for low-intensity interventions for depression in this guideline and in non-pregnant populations, taking into account the increased risk of anxiety disorders in pregnancy and the potential for harm on the fetus associated with maternal anxiety. Based on expert consensus judgement, the GDG recommended that facilitated self-help be considered for women with persistent subthreshold symptoms of anxiety in pregnancy or the postnatal period, and delivered consistently with facilitated self-help recommended for persistent subthreshold symptoms of depression, namely using CBT-based self-help materials over two to three months and being supported in using the materials (either face-to-face or by telephone) for a total of two to three hours over six sessions. In the absence of evidence for the treatment of an axiety disorder in pregnancy, the GDG considered it reasonable to extrapolate from a non-pregnant population, and recommended that low-intensity or high-intensity psychological interventions be offered in line with recommendations as set out in the NICE guidelines for *Generalised Anxiety Disorder and Panic Disorder* (NICE, 2011a), *Obsessive-Compulsive Disorder* (NICE, 2005b), *Social Anxiety Disorder* (NICE, 2013a) and *Post-Traumatic Stress Disorder* (NICE, 2005a). The GDG considered it important to highlight that only high-intensity psychological interventions are recommended for PTSD, high-intensity psychological interventions are the first-line treatment for social anxiety disorder, and healthcare professionals should be aware that progress needs to be closely monitored and stepped up to a high-intensity psychological intervention within two weeks.

There was no evidence for the treatment of severe mental illness (psychosis, schizophrenia and bipolar disorder) in pregnancy or the postnatal period, and the GDG considered that a psychological intervention in line with the guidelines on *Psychosis and Schizophrenia in Adults* (NICE, 2014) and *Bipolar Disorder* (NICE, 2006) should be considered, particularly for women who have stopped taking

psychotropic medication when they find out they are pregnant, or are changing their medication to one with a lower risk profile.

There was no evidence for the treatment of eating disorders in pregnancy or the postnatal period, and the GDG considered that a psychological intervention in line with the guideline on eating disorders (NICE, 2004a) should be offered. The GDG were, however, concerned about the potential for misinterpretation of advice that it is not necessary 'to eat for two' as validation for continuing with restrictive calorie intake or purging and the GDG recommended, based on consensus judgement and clinical opinion, that the importance of healthy eating during pregnancy and the postnatal period should be discussed, and the woman's condition should be monitored carefully throughout pregnancy and the postnatal period, including assessing the need for a fetal growth scan. The GDG also recommended that women with eating disorders in the postnatal period should be advised about, and supported in, feeding their baby, based on consensus opinion and the findings of the qualitative review of experience of care (see Chapter 6), where the need for individualized infant feeding advice for women with eating disorders emerged as a theme.

There was low quality, single study evidence for large effects associated with post-traumatic birth counselling for women who had experienced a physically traumatic birth on depression and anxiety symptomatology. However, there was also evidence for harms associated with post-traumatic birth counselling with a large effect favouring treatment as usual for a continuous measure of feelings of self-blame. These inconsistent effects may be indicative of the need for individualized information and support following a miscarriage or an obstetrically traumatic birth and this was also a theme which emerged from the qualitative review of service user experience (Chapter 6). Thematic analysis of post-traumatic birth experiences also highlighted benefits of partner involvement in discussion and debriefing (Chapter 6). Based on the quantitative and qualitative evidence, and GDG consensus opinion, the GDG recommended that women who have had a traumatic birth or miscarriage and wish to talk about their experience should be offered advice and support, and the effect of the birth or miscarriage on the partner should be taken into account.

The evidence for protocols associated with stillbirth was inconclusive with data suggestive of both benefits and harms. Data from one nested cohort study suggested that there may be harms associated with seeing and/or holding the stillborn infant, conversely findings from two cohort studies imply that there may be benefits associated with spending as much time with the stillborn infant as women wished or holding the stillborn infant. These equivocal findings are also observed in the qualitative review of service user experience (Chapter 6) where mixed opinions and experiences of photographs and mementoes following termination of a pregnancy because of fetal abnormality highlight the importance of individualised treatment. The mixed evidence, importance of individual choice and potential for harm led the GDG to consider protocols following stillbirth as a key priority for implementation and recommended that women together with their partner and family should be

offered the option of seeing a photograph of the baby, keeping mementoes of the baby such as handprints or footprints, and seeing and/or holding the baby, and should have the opportunity to discuss these options and be supported in their decision making. This should be facilitated by an experienced healthcare professional and the woman and her partner and family should be offered a follow-up appointment in primary or secondary care. The GDG were also mindful that planning should be incorporated into care plans and discussions for women who know in advance during pregnancy that their baby has died in utero. This was also an emerging theme in the qualitative review of service user experience (Chapter 6) where women expressed a desire to be provided with information and support to prepare them for making a decision about whether to see and/or hold the dead baby and for decisions about a funeral.

The GDG recognised that mental health problems may affect the mother-baby relationship and that women may be reluctant to disclose problems, and in light of potentially important safeguarding issues, recommended that assessment and monitoring of the mother–infant relationship (including verbal interaction, emotional sensitivity and physical care) should be a part of all postnatal assessments, including discussion of any concerns that the woman has about her relationship with her baby, and provision of information and treatment for the mental health problem. The evidence for interventions which directly targeted the mother–infant relationship was mixed, but largely non-significant. This inconclusive evidence prompted the GDG to recommend a definitive trial of a mother–infant relationship intervention that examines clinical and cost effectiveness and reports on the mental health of the woman, the emotional and cognitive development of the baby, and the quality of the interaction with a follow-up period of at least two years. The GDG were mindful that for some women problems in the mother-baby relationship may resolve with effective treatment of the mental health problem and for these women reassurance may be important. There was some evidence (of high to low quality from up to two studies) that treating the depression with structured psychological interventions (CBT or IPT) may have indirect statistically and clinically meaningful benefits on mother–infant attachment and there was some evidence that benefits may be maintained at long-term follow-up. However, the GDG were also aware that problems in the mother-baby relationship may not always or automatically resolve following intervention targeted at the mental health problem, and for these women further intervention for the mother-baby relationship should be considered.

7.7 RECOMMENDATIONS

7.7.1 Clinical recommendations

Using and modifying NICE guidelines for specific mental health problems

Assessment and treatment in pregnancy and the postnatal period

7.7.1.1 Use this guideline in conjunction with the NICE guideline for a specific mental health problem (see the related NICE guidance in section 3.2 [in the NICE guideline]) to inform assessment and treatment decisions in pregnancy and the postnatal period, and take into account:

- any variations in the nature and presentation of the mental health problem in pregnancy or the postnatal period
- the setting for assessment and treatment (for example, primary or secondary care services or in the community, the home or remotely by phone or computer)
- recommendations 5.4.8.5 5.4.8.5 to 5.4.8.10 in this guideline on assessment in pregnancy and the postnatal period
- recommendations 8.9.1.6 to 8.9.1.33 in this guideline on starting, using and stopping treatment in pregnancy and the postnatal period
- recommendations 5.4.8.13 to 5.4.8.14, 7.7.1.6 to 7.7.1.17 and 8.9.1.35 to 8.9.1.48 in this guideline on treating specific mental health problems in pregnancy and the postnatal period. **[new 2014]**

Treatment decisions, advice and monitoring for women who are planning a pregnancy, pregnant or in the postnatal period

Starting, using and stopping treatment

General advice

7.7.1.2 Before starting any treatment in pregnancy and the postnatal period, discuss with the woman the higher threshold for pharmacological interventions arising from the changing risk–benefit ratio for psychotropic medication at this time and the likely benefits of a psychological intervention. **[new 2014]**

7.7.1.3 If the optimal treatment for a woman with a mental health problem is psychotropic medication combined with a psychological intervention, but she declines or stops taking psychotropic medication in pregnancy or the postnatal period, ensure that:

- she is adequately supported and
- has the opportunity to discuss the risk associated with stopping psychotropic medication and
- is offered, or can continue with, a psychological intervention. **[new 2014]**

Providing interventions in pregnancy and the postnatal period

7.7.1.4 All interventions for mental health problems in pregnancy and the postnatal period should be delivered by competent practitioners. Psychological and psychosocial interventions should be based on the relevant treatment manual(s), which should guide the structure and duration of the intervention. Practitioners should consider using competence frameworks developed from the relevant treatment manual(s) and for all interventions practitioners should:

- receive regular high-quality supervision
- use routine outcome measures and ensure that the woman is involved in reviewing the efficacy of the treatment
- engage in monitoring and evaluation of treatment adherence and practitioner competence – for example, by using video and audio tapes, and external audit and scrutiny where appropriate. **[new 2014]**[12]

7.7.1.5 When a woman with a known or suspected mental health problem is referred in pregnancy or the postnatal period, assess for treatment within 2 weeks of referral and provide psychological interventions within 1 month of initial assessment. **[new 2014]**

[12] Adapted from the guideline on depression in adults (NICE clinical guideline 90).

Treating specific mental health problems in pregnancy and the postnatal period

Interventions for depression

7.7.1.6 For a woman with persistent subthreshold depressive symptoms, or mild to moderate depression, in pregnancy or the postnatal period, consider facilitated self-help (delivered as described in recommendation 1.4.2.2 of the guideline on depression in adults [NICE clinical guideline 90]). **[new 2014]**

7.7.1.7 For a woman with moderate or severe depression in pregnancy or the postnatal period, consider the following options:

- a high-intensity psychological intervention (for example, CBT)
- a TCA, SSRI or (S)NRI if the woman understands the risks associated with the medication and the mental health problem in pregnancy and the postnatal period and:
  - she has expressed a preference for medication, or
  - she declines psychological interventions, or
  - her symptoms have not responded to psychological interventions,
- a high-intensity psychological intervention in combination with medication if the woman understands the risks associated with the medication and the mental health problem in pregnancy and the postnatal period and there is no response, or a limited response, to a high-intensity psychological intervention or medication alone. **[new 2014]**

7.7.1.8 If a woman who is taking a TCA, SSRI or (S)NRI for mild to moderate depression becomes pregnant, discuss stopping the medication gradually and consider facilitated self-help (delivered as described in recommendation 1.4.2.2 of the guideline on depression in adults [NICE clinical guide 90]). **[new 2014]**

7.7.1.9 If a pregnant woman is taking a TCA, SSRI or (S)NRI for moderate depression and wants to stop her medication, take into account previous response to treatment, stage of pregnancy, risk of relapse, risk associated with medication and her preference, and discuss with her the following options:

- switching to a high-intensity psychological intervention (for example, CBT)
- changing medication if there is a drug that is effective for her with a lower risk of adverse effects. **[new 2014]**

7.7.1.10 If a pregnant woman is taking a TCA, SSRI or (S)NRI for severe depression, take into account previous response to treatment, stage of pregnancy, risk of relapse, risk associated with medication and her preference, and discuss with her the following options:

- continuing with the current medication

- changing medication if there is a drug that is effective for her with a lower risk of adverse effects
- combining medication with a high-intensity psychological intervention (for example, CBT)
- switching to a high-intensity psychological intervention (for example, CBT) if she decides to stop taking medication. **[new 2014]**

Interventions for anxiety disorders

7.7.1.11 For a woman with tokophobia (an extreme fear of childbirth), offer an opportunity to discuss her fears with a healthcare professional with expertise in providing perinatal mental health support in line with section 1.2.9 of the guideline on caesarean section (NICE clinical guideline 132). **[new 2014]**

7.7.1.12 For a woman with persistent subthreshold symptoms of anxiety in pregnancy or the postnatal period, consider facilitated self-help. This should consist of use of CBT-based self-help materials over 2–3 months with support (either face to face or by telephone) for a total of 2–3 hours over 6 sessions. **[new 2014]**

7.7.1.13 For a woman with an anxiety disorder in pregnancy or the postnatal period, offer a low-intensity psychological intervention (for example, facilitated self-help) or a high-intensity psychological intervention (for example, CBT) as initial treatment in line with the recommendations set out in the NICE guideline for the specific mental health problem and be aware that:

- only high-intensity psychological interventions are recommended for post-traumatic stress disorder
- high-intensity psychological interventions are the initial treatment for social anxiety disorder
- • progress should be closely monitored and a high-intensity psychological intervention offered within 2 weeks if symptoms have not improved. **[new 2014]**

7.7.1.14 If a woman who is taking a TCA, SSRI or (S)NRI for an anxiety disorder becomes pregnant, discuss with her the following options:

- stopping the medication gradually and switching to a high-intensity psychological intervention (for example, CBT)
- continuing with medication if she understands the risks associated with the medication and the mental health problem in pregnancy and the postnatal period and:
  - has expressed a preference for medication, **or**
  - declines psychological interventions, **or**
  - her symptoms have not responded to psychological interventions.
- changing medication if there is a drug that is effective for her with a lower risk of adverse effects

- combining medication with a high-intensity psychological intervention (for example, CBT) if the woman understands the risks associated with the medication and the mental health problem in pregnancy and the postnatal period and there is no response, or a limited response, to a high-intensity psychological intervention alone. **[new 2014]**

Psychological interventions for eating disorders

7.7.1.15 For a woman with an eating disorder in pregnancy or the postnatal period:

- offer a psychological intervention in line with the guideline on eating disorders (NICE clinical guideline 9)
- monitor the woman's condition carefully throughout pregnancy and the postnatal period
- assess the need for fetal growth scans
- discuss the importance of healthy eating during pregnancy and the postnatal period in line with the guideline on maternal and child nutrition (NICE public health guidance 11)
- advise her about feeding the baby in line with the guideline on maternal and child nutrition (NICE public health guidance 11) and support her with this. **[new 2014]**

Interventions for severe mental illness

7.7.1.16 Consider psychological interventions for women with bipolar disorder. This includes:

- CBT, IPT and behavioural couples therapy for bipolar depression
- structured individual, group and family interventions designed for bipolar disorder to reduce the risk of relapse, particularly when medication is changed or stopped. **[new 2014]**

7.7.1.17 Consider psychological interventions (CBT or family intervention) delivered as described in section 1.3.7 of the guideline on psychosis and schizophrenia in adults (NICE clinical guideline 178) for a woman with psychosis or schizophrenia who becomes pregnant and is at risk of relapse arising from:

- stress associated with pregnancy or the postnatal period
- a change in medication, including stopping antipsychotic medication. **[new 2014]**

Considerations for women and their babies in the postnatal period

Traumatic birth, stillbirth and miscarriage

7.7.1.18 Offer advice and support to women who have had a traumatic birth or miscarriage and wish to talk about their experience. Take into account the effect of the birth or miscarriage on the partner and encourage them to accept support from family and friends. **[new 2014]**

7.7.1.19 Offer women who have post-traumatic stress disorder, which has resulted from a traumatic birth, miscarriage, stillbirth or neonatal death, a high-intensity psychological intervention (trauma-focused CBT or eye movement desensitisation and reprocessing [EMDR]) in line with the guideline on post-traumatic stress disorder (PTSD) (NICE clinical guideline 26). **[new 2014]**

7.7.1.20 Do not offer single-session high-intensity psychological interventions with an explicit focus on 're-living' the trauma to women who have a traumatic birth. [**new 2014**]

7.7.1.21 Discuss with a woman whose baby is stillborn or dies soon after birth, and her partner and family, the option of 1 or more of the following:

- seeing a photograph of the baby
- having mementos of the baby
- seeing the baby
- holding the baby.

This should be facilitated by an experienced practitioner and the woman and her partner and family should be offered a follow-up appointment in primary or secondary care. If it is known that the baby has died in utero, this discussion should take place before the delivery, and continue after delivery if needed. **[new 2014]**

The mother–baby relationship

7.7.1.22 Recognise that some women with a mental health problem may experience difficulties with the mother–baby relationship. Assess the nature of this relationship, including verbal interaction, emotional sensitivity and physical care, at all postnatal contacts. Discuss any concerns that the woman has about her relationship with her baby and provide information and treatment for the mental health problem. **[new 2014]**

7.7.1.23 Consider further intervention to improve the mother–baby relationship if any problems in the relationship have not resolved. **[new 2014]**

7.7.2 Research Recommendations

7.7.2.1 What methods can improve the identification of women at high risk of postpartum psychosis and reduce this risk?

7.7.2.2 Are interventions designed to improve the quality of the mother–baby relationship in the first year after childbirth effective in women with a diagnosed mental health problem?

7.7.2.3 Is structured clinical management for moderate to severe personality disorders in pregnancy and the postnatal period effective at improving outcomes for women and their babies?

7.7.2.4 Are psychological interventions effective for treating moderate to severe anxiety disorders (including OCD, panic disorder, post-traumatic stress disorder and social anxiety disorder) in pregnancy?

7.7.2.5 What screening tools are effective in identifying the range of eating disorders (including anorexia nervosa, bulimia, binge eating disorder and eating disordrs not otherwise specified) in pregnancy?

7.7.2.6 What adaptations to current effective psychological interventions (for example, mode of delivery, duration, content, and intensity of treatment) are needed for use in the perinatal period to treat eating disorders?

8 PHARMACOLOGICAL AND PHYSICAL INTERVENTIONS

8.1 INTRODUCTION

Decisions about the use of psychotropic medication during pregnancy and in breastfeeding are difficult, both for women with psychiatric illness and for the clinicians who look after them. In making these decisions, the risks and benefits of all options must be considered, taking into account a woman's individual history and circumstances. A range of management approaches may be appropriate including improved support and specific psychological or social interventions but for many women treatment with medication will be an important therapeutic option.

It is important to recognise that there are many different scenarios in which women may be prescribed psychotropic medication in the perinatal period. These include the new onset of an episode of psychiatric disorder, which may be the first episode or a recurrence of a previous diagnosis, or the prophylaxis of pre-existing illness in women who are currently well. Each of these particular situations raises specific issues and may lead to different decisions about particular medication that may be chosen.

There are a number of reasons why merely reporting a reproductive safety league table for each medication class is problematic. Each woman is an individual with her own history of illness and previous response to medication. For this reason, for many scenarios there are no clear right and wrong answers, and in this chapter we go further than merely reporting the relevant studies on reproductive safety and discuss the general principles of managing women with psychotropic medication in pregnancy and breastfeeding.

In weighing up the risks and benefits of using medication in the perinatal period an important consideration is the increased risk of severe episodes of mental illness in relation to childbirth. For some women, those with a previous severe postpartum episode or an existing diagnosis of bipolar disorder for example, the immediate postpartum is a period of very high risk and decisions about medications must be made against this background. It is also important to recognise that episodes of severe psychiatric illness may have negative consequences for the woman, her baby and her family, and these must be weighed against what is known about the risks of taking medication.

Any increased risk associated with the use of medication must be interpreted against the background malformation rate in the general population of between 2 and 4%. In addition, when considering the reproductive safety of psychotropic medication, it is important to go beyond merely teratogenic risks and also consider issues of neonatal withdrawal and of longer term effects on cognitive development or behaviour. In

this regard, it is important to consider the particular stage of pregnancy as risks may differ considerably in each trimester.

As will become clear through this chapter, the amount of data we have varies hugely between and even within medication class. For some medications we have data on tens of thousands of pregnancy exposures, for others we may have a few case reports or even no data at all. It is vital, therefore, that we do not interpret the lack of evidence of harm as evidence of safety. For some medications, even for those such as lithium that have been used for many decades, our evidence base may be very limited. For other medications, antiepileptic medications for example, although the evidence base is larger, it may come from the treatment of other conditions, with little data on use in psychiatric disorders.

However, it is important to note that the use of data from an indirect population (women with epilepsy) does not necessarily invalidate the evidence, which was still seen as relevant to women with bipolar disorder. Moreover, the larger dataset and the small number of anticonvulsant drugs used in bipolar disorder may enable consideration of individual drugs which is important where there are grounds to believe that the safety profiles may be different for different drugs within a class. Even where there are extensive data, such as is the case with SSRI antidepressants, it remains difficult to know whether any increase in risk that has been identified is due to the medication being taken, to the underlying psychiatric disorder itself, to an overlap in genetic vulnerability or to other factors associated with psychiatric disorders and the use of medication.

In helping women through these difficult decisions, clinicians must help women to weigh up the risks and benefits of all options in the context of their individual history and circumstances. Although the communication of risk is a vital and difficult area of clinical practice and an emerging area of research, more research is clearly needed to address the particular issues around discussing psychotropic medication in pregnancy with women and their partners.

This chapter is divided into eight main sections, comprising six reviews. Section 8.2 reviews the evidence for pharmacological interventions for the prevention of mental health problems in pregnancy and the postnatal period – the review is separated into evidence for the effects on outcomes for women with no identified risk factors, on outcomes for women with identified risk factors, and on the prophylaxis of mental health problems. Section 8.3 reviews the evidence for the efficacy of pharmacological interventions for the treatment of mental health problems in pregnancy and the postnatal period. Section 8.4 reviews the harms associated with specific types of drugs in pregnancy and the postnatal period, including antidepressants, antipsychotics, anticonvulsants, lithium, benzodiazepines and stimulants. Sections 8.5 and 8.6 review physical interventions for the prevention of mental health problems in pregnancy and the postnatal period and their treatment, respectively. These interventions include physical activity, acupuncture, massage and bright light therapy. Section 8.7 comprises a separate review of electroconvulsive therapy (ECT).

Because of the need to balance the risks and benefits of treatment in pregnancy and the postnatal period, the GDG wished to consider the evidence for pharmacological and physical interventions as a whole; therefore all of their decisions are set out in Section 8.8, rather than after each individual review. The recommendations themselves follow in Section 8.9.

8.2 PHARMACOLOGICAL INTERVENTIONS FOR THE PREVENTION OF MENTAL HEALTH PROBLEMS IN PREGNANCY AND THE POSTNATAL PERIOD

8.2.1 Clinical review protocol (prevention)

The review protocol summary, including the review question(s), information about the databases searched, and the eligibility criteria used for this section of the guideline, can be found in Table 263. A complete list of review questions can be found in Appendix 8; further information about the search strategy can be found in Appendix 10; the full review protocols can be found in Appendix 9.
The review strategy was to evaluate the clinical effectiveness of the pharmacological interventions using meta-analysis. However, in the absence of adequate data, the available evidence was synthesised using narrative methods. An analysis of all interventions was conducted and graded.

Table 263: Clinical review protocol summary for the review of pharmacological interventions for the prevention of mental health problems

Component	Description
Review question(s)	RQ 2.1 What is the effectiveness of selective preventative interventions (for women with no risk factors) in reducing the likelihood of developing mental health problems in pregnancy or the postnatal period? RQ 2.2 What is the effectiveness of indicated preventative interventions (for women with identified risk factors present) in reducing the likelihood of developing mental health problems in pregnancy or the postnatal period? RQ 2.3 What strategies should be adopted to minimise potential harm to the women or the fetus/infant of these interventions?
Population	**Included** **Review question 2.1** Women who are pregnant or postnatal (from delivery to the end of the first year). Inclusion is not based on any other baseline risk factors. **Review question 2.2** Women who are pregnant or postnatal (from delivery to the end of the first year) who are considered to be 'at risk' of developing mental health problems. Include women: • with a history of a mental health problem but who do not meet diagnostic criteria for mental health problems at the current time • experiencing major life events • with a family history of mental health problems

	• with psychosocial risk factors (for example socioeconomic status) • who have infants with regulatory problems • who experienced an operative delivery or traumatic birth • who experienced a pre-term delivery (<37 weeks gestation) and/or whose infant had a low birth weight • who experienced a miscarriage • who are adolescents • experiencing Intimate Partner Violence Exclude women: who are currently receiving treatment (psychosocial or pharmacological) for an existing mental health problem (see review of interventions for the treatment of a mental health problem) who are not pregnant or postnatal period (up to 1 year postnatal)
Intervention(s)	**Included interventions** Pharmacological interventions for women with no pre-specified baseline risk factors (other than being pregnant or in the postnatal period) (RQ 2.1) or for women with at least one identified baseline risk factor (RQ 2.2), including: • Psychotropic medications • Dietary supplements • Hormones **Excluded Interventions** • Universal prevention programmes (that is, targeted to the general public or to a whole population group that has not been identified on the basis of increased risk)
Comparison	Review question 2.1 and 2.2 Treatment as usual, enhanced treatment as usual, no treatment, waitlist control Another active prevention intervention
Critical outcomes	• Maternal Outcomes • Symptom-based • Diagnosis of mental health problem • Symptomatology (clinician- and self-report) • Relapse • Service utilisation - Hospitalisation for mental health problems - Retention in services (assessed through dropout rates as a proxy measure) • Experience of care - Satisfaction - Acceptability of treatment (including dropout as a proxy measure) • Quality of life - Quality of life measures - Functional disability - Social functioning - Social support - Perceived parenting stress • Harm - Side effects (including dropout because of side effects) • Quality of mother-infant interaction and infant care - Quality of mother-infant interaction measures (including maternal sensitivity and child responsivity)

	- Establishing or continuing breastfeeding **Fetal/Infant outcomes** • Fetal and infant physical development (including congenital malformations) • Side effects • Cognitive development of the infant • Physical development of the infant • Emotional development of the infant • Optimal care of infant (for example vaccinations, well-baby check-ups) • Prevention of neglect or abuse of the infant • Service use - Planned (health visitor, vaccinations, well-baby check-ups) - Unplanned (emergency department visits, inpatient, urgent or acute care) - Social service involvement - Mode of delivery
Study design	Review question 2.1 and 2.2 Systematic reviews of RCTs Primary RCTs Review question 2.3 N/A; GDG consensus-based
Note.	

8.2.2 Studies considered[13]

Women with no identified risk factors

Four RCTs met the eligibility criteria for this review: HARRISONHOHNER2001 (Harrison-Hohner et al., 2001), LLORENTE2003 (Llorente et al., 2003), MAKRIDES2010 (Makrides et al., 2010), MOKHBER2011 (Mokhber et al., 2011). All of these studies were published in peer-reviewed journals between 2001 and 2011. Further information about the included studies can be found in Appendix 18.
All studies included sufficient data to be included in the statistical analysis. Of these, there were two studies (N=2,537) involving a comparison of omega-3 and placebo, one study (N=166) that compared selenium and placebo and one study (N=374) that compared calcium and placebo (see Table 264).

Women with identified risk factors

Two RCTs met the eligibility criteria for this review: HARRIS2002 (Harris et al., 2002), LAWRIE1998B (Lawrie et al., 1998b). Further information about included studies can be found in Appendix 18.

There was one study (N=180) that compared norethisterone with placebo and one study (N=446) involved a comparison between thyroxine and placebo (see Table

[13] Here and elsewhere in the guideline, each study considered for review is referred to by a study ID in capital letters (primary author and date of study publication, except where a study is in press or only submitted for publication, then a date is not used).

265). In one study participants had psychosocial risk factors (low income) and in one study participants were positive for thyroid antibodies (although this was not one of the pre-specified risk factors, women included in this study were at risk of postnatal depression, and therefore included in the review for risk factors).

Prophylaxis of mental health problems

Two RCTs met the eligibility criteria for this review: WISNER2001 (Wisner et al., 2001); WISNER2004B (Wisner et al., 2004b). In addition seven studies were excluded from the review. The reasons for exclusion were that the studies did not have a control group or were not RCTs. Further information about both included and excluded studies can be found in Appendix 18. One study compared TCAs (nortriptyline) with placebo, and one study compared SSRIs (sertraline) with placebo (see

Table 266).

Table 264: Study information table for trials included in the meta-analyses of pharmacological interventions compared with placebo in women with no identified risk factors

	Omega-3 versus placebo	Selenium versus placebo	Calcium versus placebo
Total no. of trials (k); participants (N)	2 (2,537)	1 (166)	1 (374)
Study ID	LLORENTE2003 MAKRIDES2010	MOKHBER2011	HARRISONHOHNER2001[2]
Country	(1) US (2) Australia	Iran	US
Mean Age of Paricipants (years)	(1) 31 (2) 29	22	22
Timing of intervention	(1) Postnatal (2) Pregnancy	Pregnancy	Pregnancy
Dose (mean)	(1) 200 mg DHA/ day (2) Three 500-mg DHA/ day	100 mg/ day	2,000 mg/Taken in split dose (morning and evening meals)
Length of intervention (weeks)	(1) 17 (2) Approx: 19 (22 weeks gestation to birth)	Approx: 26 (first trimester of pregnancy until delivery)	Approx: 19 (13-21 weeks through to delivery)
Time points[1]	(1) Post-treatment; long-term follow-up (2) Post-treatment; Intermediate follow-up	Post-treatment	Post-treatment; short-term follow-up
Setting	(1) Clinic (primary) (2) Clinic (primary)	Clinic (primary)	Clinic (primary)
Intervention	(1)-(2) Omega-3 (DHA)	Selenium	Elemental calcium
Comparison	(1) Identical capsules (2) Vegetable oil capsules	Matching yeast tablets	Tablets identical to calcium tablets

Note. Abbreviations: DHA=Docosahaxaenoic acid
[1]Time points: Post-treatment or first measurement; Short-term follow-up (9-16 weeks post-intervention); Intermediate follow-up (17-24 weeks post-intervention); Long-term follow-up (25-103 weeks post-intervention follow-up); Very long-term follow-up (≥104 weeks).
[2] Participants recruited from cohort from previous ongoing trial LEVINE1997 (Levine et al.; 1997)

Table 265: Study information table for trials included in the meta-analyses of any pharmacological intervention versus placebo comparison in women with identified risk factors

	Norethisterone versus placebo	**Thyroxine versus placebo**
Total no. of trials (k); participants (N)	1 (180)	1 (446)
Study ID	LAWRIE1998B	HARRIS2002
Country	South Africa	UK
Mean Age of Paricipants (years)	32	29
Timing of intervention	Postnatal	Postnatal
Mean dose	200 mg	100 mg/ day
Length of intervention	Single dose within 48 hours of delivery	20 weeks
Risk factor	Low income urban population	Women positive for thyroid antibodies in early gestation[2]
Time points[1]	Post-treatment; Short-term follow-up	Post-treatment
Setting	Clinic (primary)	Clinic (primary)
Intervention	Norethisterone	Thyroxine
Comparison	Placebo	Placebo tablet

[1]Time points: Post-treatment or first measurement; Short-term follow-up (9-16 weeks post-intervention); Intermediate follow-up (17-24 weeks post-intervention); Long-term follow-up (25-103 weeks post-intervention follow-up); Very long-term follow-up (≥104 weeks).
[2] Although this was not one of the pre-specified risk factors, women included in this study were a'prone to postnatal depression', and therefore included in the review for risk factors

Table 266: Study information table for trials included in the meta-analyses of any pharmacological intervention versus placebo comparison for prophylaxis of mental health problems

	TCA (nortriptyline) versus placebo	SSRI (sertraline) versus placebo
Total no. of trials (k); participants (N)	1 (56)	1 (25)
Study ID	WISNER2001	WISNER2004B
Country	US	US
Mean age of paricipants (years)	NR	32
Timing of intervention	Postnatal	Postnatal
Mean dose	20-70 mg increased and tapered	25-75 mg increased and tapered
Length of intervention (weeks)	20	17
Risk factor	At least one past episode of postnatal major depression	At least one past episode of postnatal major depression
Time points[1]	Post-treatment; intermediate follow-up (26 weeks)	Post-treatment
Setting	Clinic (primary)	Clinic (primary)
Intervention	TCAs (nortriptyline)	SSRIs (sertaline)
Comparison	Placebo	Placebo

Note.
[1] Time points: post-treatment or first measurement; short-term follow-up (9-16 weeks post-intervention); intermediate follow-up (17-24 weeks post-intervention); long-term follow-up (25-103 weeks post-intervention follow-up); very long-term follow-up (≥104 weeks).

8.2.3 Clinical evidence for preventative effects on outcomes for women with no identified risk factors

Summary of findings can be found in the tables presented in this section. The full GRADE evidence profiles and associated forest plots can be found in Appendix 22 and Appendix 19, respectively.

Depression outcomes (by intervention)

Omega-3 versus placebo

There was no evidence for clinically or statistically significant benefits (p=0.18–1.00) associated with omega-3 for mean depression scores, depression symptomology or diagnosis at endpoint or at intermediate follow-up (Table 267).

Table 267: Summary of findings table for effects of omega-3 compared with placebo on preventing depression outcomes in women with no identified risk factors

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Depression: Omega-3 versus placebo				
Depression mean scores (Post-treatment) BDI Follow-up: mean 17 weeks		The mean depression mean scores (post-treatment) in the intervention groups was 0.15 standard deviations higher (0.26 lower to 0.57 higher)		89 (1 study)	⊕⊖⊖⊖ very low1,2	SMD 0.15 (-0.26 to 0.57)
Depression mean scores (Long-term follow-up, 25-103 weeks post-intervention) EPDS Follow-up: mean 61 weeks		The mean depression mean scores (long-term follow-up, 25-103 weeks post-intervention) in the intervention groups was 0 standard deviations higher (0.49 lower to 0.49 higher)		63 (1 study)	⊕⊕⊖⊖ low1,3	SMD 0 (-0.49 to 0.49)
Depression symptomology (Post treatment) EPDS >12 Follow-up: mean19 weeks	Study population 109 per 1000	96 per 1000 (76 to 122)	RR 0.88 (0.7 to 1.12)	2399 (1 study)	⊕⊕⊕⊖ moderate4	
	Moderate 109 per 1000	96 per 1000 (76 to 122)				
Depression symptomology (Intermediate follow-up, 17-24 weeks post-intervention) EPDS >12 Follow-up: mean 24 weeks	Study population 115 per 1000	98 per 1000 (77 to 123)	RR 0.85 (0.67 to 1.07)	2399 (1 study)	⊕⊕⊕⊖ moderate4	
	Moderate 115 per 1000	98 per 1000 (77 to 123)				
Depression diagnosis (current depression new or existing during study	Study population 55 per 1000	51 per 1000 (37 to 72)	RR 0.93 (0.67 to 1.31)	2448 (2 studies)	⊕⊕⊕⊖ moderate4	

period) SCID diagnosis/unknown diagnostic test Follow-up: mean 17-19 weeks	Moderate 95 per 1000	88 per 1000 (64 to 124)

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Risk of bias due to high attrition
2 Total population size is less than 400 (a threshold rule-of-thumb) and 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
3 Total population size is less than 400 (a threshold rule-of-thumb)
4 Total number of events is less than 300 (a threshold rule-of-thumb)

Selenium versus placebo

There was low quality, single study (N=85) evidence in favour of a preventative benefit of selenium on reducing mean depression scores at endpoint, however this effect did not reach statistical significance (p=0.07) and failed to reach a threshold indicative of clinically significant benefits (Table 268).

Table 268: Summary of findings table for effects of selenium compared with placebo on preventing depression outcomes in women with no identified risk factors

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Depression: Selenium versus placebo	Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
Depression mean scores (Post-treatment) EPDS Follow-up: 8 weeks		The mean depression mean scores (post-treatment) in the intervention groups was 0.39 standard deviations lower		85 (1 study)	⊕⊕⊖⊖ low1,2	SMD -0.39 (-0.82 to 0.04)

	(0.82 lower to 0.04 higher)

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Unclear selection bias
2 Total population size is less than 400 (a threshold rule-of-thumb)

Calcium versus placebo

The evidence for calcium as a preventative intervention was inconsistent (Table 269). There was low quality, single study (N=374) evidence for a moderate preventative benefit of calcium on depression symptomology ay endpoint, however this effect was not statistically significant (p=0.13) and there was very serious imprecision (due to the small event rate and the 95% CIs included both no effect and appreciable benefit). There was some discrepancy between dichotomous and continuous measures of depression at short term follow-up. There was moderate quality, single study (N=247) evidence for a large beneficial effect of calcium on preventing depression symptomology (p=0.02), however there was serious imprecision of this effect estimate due to the low number of events. In addition, there was no statistically or clinically significant preventive benefit on mean depression scores at short-term follow-up (p=0.13).

Table 269: Summary of findings table for effects of calcium compared with placebo on depression outcomes in women with no identified risk factors

Depression: Calcium versus placebo for prevention (no risk factors present)

Patient or population: patients with prevention (no risk factors present)
Settings:
Intervention: Depression: Calcium versus placebo

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				

	Control	Depression: Calcium versus placebo				
Depression mean scores (Short-term follow-up, 9-16 weeks post-intervention) EPDS Follow-up: 12 weeks		The mean depression mean scores (short-term follow-up, 9-16 weeks post-intervention) in the intervention groups was 0.19 standard deviations lower (0.44 lower to 0.06 higher)		247 (1 study)	⊕⊕⊕⊖ moderate1	SMD -0.19 (-0.44 to 0.06)
Depression symptomology (Post-treatment) EPDS >=14 Follow-up: 6 weeks	Study population 187 per 1000 Moderate 187 per 1000	135 per 1000 (84 to 217) 135 per 1000 (84 to 217)	RR 0.72 (0.45 to 1.16)	374 (1 study)	⊕⊕⊖⊖ low2	
Depression symptomology (Short-term follow-up, 9-16 weeks post-intervention) EPDS >=14 Follow-up: 12 weeks	Study population 153 per 1000 Moderate 153 per 1000	57 per 1000 (25 to 130) 57 per 1000 (24 to 130)	RR 0.37 (0.16 to 0.85)	247 (1 study)	⊕⊕⊕⊖ moderate1	

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total population size is less than 400 (a threshold rule-of-thumb)
2 Total number of events is less than 300 (a threshold rule-of-thumb) and 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Compliance outcomes (by intervention)

Selenium versus placebo

There was low quality, single study (N=85) evidence for a large beneficial effect of selenium on compliance (as measured by serum selenium concentration)

post-treatment (p<0.00001, Table 270). However, confidence that this is a true measure of the effect is low due to the small population size and unclear risk of selection bias (unclear method of randomisation and allocation concealment).

Table 270: Summary of findings table for effects of calcium compared with placebo on compliance outcomes in women with no identified risk factors

Compliance: Selenium versus placebo for prevention (no risk factors present)

Patient or population: patients with prevention (no risk factors present)
Settings:
Intervention: Compliance: Selenium versus placebo

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk Control	Corresponding risk Compliance: Selenium versus placebo				
Serum selenium concentration (Post-treatment) Follow-up: 26 weeks		The mean serum selenium concentration (post-treatment) in the intervention groups was 1.39 standard deviations lower (1.87 to 0.92 lower)		85 (1 study)	⊕⊕⊖⊖ **low**[1,2]	SMD -1.39 (-1.87 to -0.92)

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Unclear randomisation method or allocation concealment
[2] Total population size is less than 400 (a threshold rule-of-thumb)

Quality of life outcomes (by intervention)

Calcium versus placebo

There was no statistically or clinically significant benefit of calcium on positive (p=0.16) or negative (p=0.48) life events (Table 271).

Table 271: Summary of findings table for effects of calcium compared with placebo on quality of life outcomes in women with no identified risk factors

Quality of life: Calcium versus placebo for prevention (no risk factors present)

Patient or population: patients with prevention (no risk factors present)
Settings:
Intervention: Quality of life: Calcium versus placebo

	Illustrative comparative risks* (95% CI) Assumed risk Control	 Corresponding risk Quality of life: Calcium versus placebo	Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
Positive life events (Post-treatment) Sarason's Life Events Survey Follow-up: 6 weeks		The mean positive life events (post-treatment) in the intervention groups was 0.18 standard deviations lower (0.43 lower to 0.07 higher)		247 (1 study)	⊕⊕⊕⊖ moderate1	SMD -0.18 (-0.43 to 0.07)
Negative life events (Post-treatment) Sarason's Life Events Survey Follow-up: 6 weeks		The mean negative life events (post-treatment) in the intervention groups was 0.09 standard deviations lower (0.34 lower to 0.16 higher)		247 (1 study)	⊕⊕⊕⊖ moderate1	SMD -0.09 (-0.34 to 0.16)

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total population size is less than 400 (a threshold rule-of-thumb)

Infant outcomes (by intervention)

Omega-3 versus placebo

There was no evidence for a statistically or clinically significant benefit of omega-3 on any of the of the Bayley scales of Infant and toddler development subscales (p=0.14-0.95) at long-term follow-up. There was moderate quality, single study (N=726) evidence for a statistically significant benefit on cognitive performance using an ITT analysis (p=0.05), however this effect was just under the threshold indicative of clinically significant benefits. There was no statistically or clinically significant effect on language performance (p=0.91, Table 272).

Table 272: Summary of findings table for preventative effects of omega-3 compared with placebo on infant outcomes in women with no identified risk factors

Infant outcomes: Omega-3 versus placebo for prevention (no risk factors present)

Patient or population: patients with prevention (no risk factors present)
Settings:
Intervention: Infant outcomes: Omega-3 versus placebo

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Infant outcomes: Omega-3 versus placebo	Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
Mean development symptomology (Long-term follow-up, 25-103 weeks post-intervention) - Cognitive standardised score ITT analysis Bayley scales of Infant and toddler development Follow-up: 78 weeks		The mean development symptomology (long-term follow-up, 25-103 weeks post-intervention) - cognitive standardised score in the intervention groups was 0.01 standard deviations higher (0.14 lower to 0.15 higher)		726 (1 study)	⊕⊕⊕⊖ moderate1	SMD 0.01 (-0.14 to 0.15)
Mean development symptomology (Long-term follow-up, 25-103 weeks post-intervention) - Language		The mean development symptomology (long-term follow-up, 25-103 weeks post-intervention) -		726 (1 study)	⊕⊕⊕⊖ moderate1	SMD -0.1 (-0.25 to 0.04)

standardised score ITT analysis Bayley scales of Infant and toddler development Follow-up: 78 weeks		language standardised score in the intervention groups was 0.1 standard deviations lower (0.25 lower to 0.04 higher)			
Mean development symptomology (Long-term follow-up, 25-103 weeks post-intervention) - Motor standardised score ITT analysis Bayley scales of Infant and toddler development Follow-up: 78 weeks		The mean development symptomology (long-term follow-up, 25-103 weeks post-intervention) - motor standardised score in the intervention groups was 0.03 standard deviations lower (0.18 lower to 0.12 higher)		726 (1 study)	⊕⊕⊕⊖ moderate1	SMD -0.03 (-0.18 to 0.12)
Mean development symptomology (Long-term follow-up, 25-103 weeks post-intervention) - Social-Emotional standardised score ITT analysis Bayley scales of Infant and toddler development Follow-up: 78 weeks		The mean development symptomology (long-term follow-up, 25-103 weeks post-intervention) - social-emotional standardised score in the intervention groups was 0.05 standard deviations lower (0.02 to 0.09 lower)		726 (1 study)	⊕⊕⊕⊖ moderate1	SMD -0.25 (-0.4 to -0.11)
Mean development symptomology (Long-term follow-up, 25-103 weeks post-intervention) - Adaptive Behavior standardised score ITT analysis Bayley scales of Infant and toddler development Follow-up: 78 weeks		The mean development symptomology (long-term follow-up, 25-103 weeks post-intervention) - adaptive behaviour standardised score in the intervention groups was 0.11 standard deviations lower (0.26 lower to 0.03 higher)		726 (1 study)	⊕⊕⊕⊖ moderate1	SMD -0.11 (-0.26 to 0.03)
Delayed cognitive performance (Long-term follow-up, 25-103 weeks post-intervention) ITT analysis Bayley scales of	Study population 64 per 1000	31 per 1000 (15 to 63)	RR 0.49 (0.24 to 0.98)	726 (1 study)	⊕⊕⊕⊖ moderate1	
	Moderate 64 per 1000	31 per 1000 (15 to 63)				

infant development, <85 Follow-up: 78 weeks					
Delayed language performance (Long-term follow-up, 25-103 weeks post-intervention) ITT analysis Bayley scales of infant development, <85 Follow-up: 78 weeks	Study population		RR 1.02 (0.74 to 1.4)	726 (1 study)	⨁⨁⊖⊖ low1,2
	173 per 1000	177 per 1000 (128 to 243)			
	Moderate				
	173 per 1000	176 per 1000 (128 to 242)			

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Unclear attrition bias for follow-up data
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Leaving the study early for any reason (by intervention)

Omega-3 versus placebo

There was no evidence for a statistically or clinically significant benefit of omega-3 on leaving the study early for any reason (p=0.25, Table 273*)*

Table 273: Summary of findings table for effects of omega-3 compared with placebo on leaving the study early in women with no identified risk factors

Leaving the study early for any reason: Omega-3 versus placebo for prevention (no risk factors present)						
Patient or population: patients with prevention (no risk factors present) **Settings:** **Intervention: Leaving the study early for any reason: Omega-3 versus placebo**						
Outcomes	**Illustrative comparative risks* (95% CI)** **Assumed risk** **Control**	**Corresponding risk** **Leaving the study early for any reason: Omega-3 versus placebo**	**Relative effect (95% CI)**	**No of Participants (studies)**	**Quality of the evidence (GRADE)**	**Comments**
Leaving study early for any reason (Post-treatment) Follow-up: 17-19 weeks	Study population		RR 0.69 (0.37 to 1.3)	2,537 (2 studies)	⊕⊕⊖⊖ low1,2	
	43 per 1000	30 per 1000 (16 to 56)				
	Moderate					
	153 per 1000	106 per 1000 (57 to 199)				

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 There was evidence of substantial heterogeneity between effect sizes
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Adverse events and service utilisation (by intervention)

Omega-3 versus placebo

There was moderate quality, single study (N=2,399) evidence for moderate beneficial effect (p=0.04) of omega-3 on preventing infant admission to neonatal intensive care (Table 274). However, the imprecision of this effect estimate was serious due to the small number of events. There were no statistically or clinically significant

differences between omega-3 and placebo on maternal hospitalisation for serious adverse events (p=1.00) or major congenital abnormalities of the infant at long-term follow-up (p=0.43).

Table 274: Summary of findings table for effects of omega-3 compared with placebo on adverse events and service utilisation

Adverse events: Omega-3 versus placebo for prevention (no risk factors present)

Patient or population: patients with prevention (no risk factors present)
Settings:
Intervention: Adverse events: Omega-3 versus placebo

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Adverse events: Omega-3 versus placebo				
Maternal hospitalisation for serious adverse events (Post-treatment) Follow-up: 19 weeks	Study population		RR 1 (0.14 to 7.12)	2,399 (1 study)	⊕⊕⊖⊖ low1	
	2 per 1000	2 per 1000 (0 to 12)				
	Moderate					
	2 per 1000	2 per 1000 (0 to 14)				
Infant admission to neonatal intensive care due to adverse events (Post-treatment) Follow-up: 19 weeks	Study population		RR 0.57 (0.34 to 0.97)	2,399 (1 study)	⊕⊕⊕⊖ moderate1,2	
	31 per 1000	18 per 1000 (10 to 30)				
	Moderate					
	31 per 1000	18 per 1000 (11 to 30)				
Major congenital abnormality of the infant (Long term follow-up, 25-103 weeks post-intervention) Follow-up: 78 weeks	Study population		RR 1.37 (0.63 to 2.97)	2,399 (1 study)	⊕⊕⊕⊖ moderate1,2	
	9 per 1000	13 per 1000 (6 to 27)				
	Moderate					
	9 per 1000	12 per 1000 (6 to 27)				

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
[2] Total number of events is less than 300 (a threshold rule-of-thumb)

8.2.4 Clinical evidence for preventative effects on outcomes for women with identified risk factors

Summary of findings can be found in the tables presented in this section. The full GRADE evidence profiles and associated forest plots can be found in Appendix 22 and Appendix 19, respectively.

Depression outcomes (by intervention)

Thyroxine versus placebo

There was no evidence for a statistically or clinically significant benefit (p=0.44-0.98) of thyroxine on depression symptomology or diagnosis at the end of intervention (Table 275).

Table 275: Summary of findings table for effects of thyroxine compared with placebo on depression outcomes in women with identified risk factors

Depression: Thyroxine versus placebo for prevention (risk factors present)

Patient or population: patients with prevention (risk factors present)
Settings:
Intervention: Depression: Thyroxine versus placebo

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Depression: Thyroxine versus placebo				
Depression diagnosis, major depression-definite and probable cases (Post-treatment) RDC Follow-up: 20 weeks	Study population		RR 0.85 (0.34 to 2.16)	341 (1 study)	⊕⊕⊖⊖ low1,2	
	54 per 1000	46 per 1000 (18 to 116)				
	Moderate					
	54 per 1000	46 per 1000 (18 to 117)				
Depression diagnosis, any depression (Post-treatment)	Study population		RR 0.81 (0.48 to 1.38)	341 (1 study)	⊕⊕⊖⊖ low1	
	156 per 1000	126 per 1000 (75 to 215)				

RDC Follow-up: 20 weeks	Moderate				
	156 per 1000	126 per 1000 (75 to 215)			
Depression symptomology (Post-treatment) EPDS >=13	Study population		RR 1.01 (0.57 to 1.79)	341 (1 study)	⊕⊕⊖⊖ low1,2
	120 per 1000	121 per 1000 (68 to 214)			
	Moderate				

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Baseline scores significantly different, unclear attrition bias
[2] 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Norethisterone compared with placebo

There was moderate quality, single study (N=163) evidence for a non-beneficial effect of norethisterone on depression outcomes at the end of intervention (Table 276). There was a statistically significant effect on mean depression scores favouring the placebo group compared to the norethisterone group (p=0.004), although this effect failed to reach a threshold indicative of clinically significant benefits. There was a moderate effect favouring placebo on depression symptomology (p=0.01), however there was serious imprecision (due to the small sample size). Moreover, this effect was not maintained at short-term follow-up, with no statistically or clinically significant difference in effect on mean depression scores or symptomology.

Table 276: Summary of findings table for effects of norethisterone compared with placebo on adverse events

Depression: Norethisterone versus placebo for prevention (risk factors present)

Patient or population: patients with prevention (risk factors present)
Settings:
Intervention: Depression: Norethisterone versus placebo

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Depression: Norethisterone versus placebo	Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
Depression mean scores (Post treatment) EPDS Follow-up: 6 weeks		The mean depression mean scores (post treatment) in the intervention groups was 0.46 standard deviations higher (0.15 to 0.77 higher)		163 (1 study)	⊕⊕⊕⊖ moderate2	SMD 0.46 (0.15 to 0.77)
Depression mean scores (Short-term follow-up, 9-16 weeks post-intervention) EPDS Follow-up: 17 weeks		The mean depression mean scores (short-term follow-up, 9-16 weeks post-intervention) in the intervention groups was 0.12 standard deviations higher (0.19 lower to 0.42 higher)		168 (1 study)	⊕⊕⊕⊖ moderate2	SMD 0.12 (-0.19 to 0.42)
Depression symptomology (Post-treatment) EPDS >11 Follow-up: 6 weeks	Study population 260 per 1000 Moderate 260 per 1000	 455 per 1000 (291 to 706) 455 per 1000 (291 to 707)	RR 1.75 (1.12 to 2.72)	163 (1 study)	⊕⊕⊕⊖ moderate2	
Depression symptomology (Short-term follow-up, 9-16 weeks post-intervention) EPDS >11 Follow-up: 6 weeks	Study population 296 per 1000 Moderate 296 per 1000	 323 per 1000 (204 to 507) 323 per 1000 (204 to 506)	RR 1.09 (0.69 to 1.71)	168 (1 study)	⊕⊕⊖⊖ low1	

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total population size is less than 400 (a threshold rule-of-thumb) and 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
2 Total population size is less than 400 (a threshold rule-of-thumb)

Compliance outcomes (by intervention)

Thyroxine versus placebo

There was no evidence for a statistically or clinically significant benefit of thyroxine on compliance post-treatment (p=0.44, Table 277).

Table 277: Summary of findings table for effects of thyroxine compared with placebo on adverse events

Compliance: Thyroxine versus placebo for prevention (no risk factors present)

Patient or population: patients with prevention (no risk factors present)
Settings:
Intervention: Compliance: Thyroxine versus placebo

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Compliance: Thyroxine versus placebo				
Numbers not compliant (Post-treatment) Follow-up: 20 weeks	Study population		RR 0.88 (0.63 to 1.22)	446 (1 study)	⊕⊕⊕⊖ moderate1	
	251 per 1000	221 per 1000 (158 to 306)				
	Moderate					
	251 per 1000	221 per 1000 (158 to 306)				

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Mother-infant interaction outcomes (by intervention)

Norethisterone versus placebo

There was no evidence for a statistically or clinically significant benefit of norethisterone on breastfeeding outcomes at the end of intervention (p=0.30) or at short-term follow-up (p=0.28, Table 278).

Table 278: Summary of findings table for effects of norethisterone compared with placebo on mother-infant interaction outcomes

Mother-infant interaction: Norethisterone versus placebo for prevention (risk factors present)

Patient or population: patients with prevention (risk factors present)
Settings:
Intervention: Mother-infant interaction: Norethisterone versus placebo

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Mother-infant interaction: Norethisterone versus placebo				
Breastfeeding-exclusive or partial (Post-treatment) Follow-up: 6 weeks	Study population		RR 0.92 (0.77 to 1.08)	166 (1 study)	⊕⊕⊕⊖ moderate1	
	800 per 1000	736 per 1000 (616 to 864)				
	Moderate					
	800 per 1000	736 per 1000 (616 to 864)				
Breastfeeding-exclusive or partial (Short term follow-up, 9-16 weeks post-intervention) Follow-up: 13 weeks	Study population		RR 0.9 (0.74 to 1.09)	168 (1 study)	⊕⊕⊕⊖ moderate1	
	753 per 1000	678 per 1000 (557 to 821)				
	Moderate					
	753 per 1000	678 per 1000 (557 to 821)				

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total population size is less than 400 (a threshold rule-of-thumb

Leaving the study early (by intervention)

Norethisterone versus placebo

There was low to moderate quality, single study (N=180) evidence for large beneficial effect of norethisterone on leaving the study early at the end of intervention (p=0.03) and short-term follow-up (p=0.09), however the imprecision of this effect estimate was serious due to the small population and the 95% CIs were wide (Table 279).

Table 279: Summary of findings table for effects of omega-3 compared with placebo on adverse events

Leaving the study early: Norethisterone versus placebo for prevention (risk factors present)

Patient or population: patients with prevention (risk factors present)
Settings:
Intervention: Leaving the study early: Norethisterone versus placebo

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Leaving the study early: Norethisterone versus placebo				
Leaving study early for any reason (Post-treatment) Follow-up: 6 weeks	Study population		RR 0.31 (0.1 to 0.91)	180 (1 study)	⊕⊕⊕⊖ moderate1	
	144 per 1000	45 per 1000 (14 to 131)				
	Moderate					
	144 per 1000	45 per 1000 (14 to 131)				
	Study population					

Leaving the study early for any reason (short-term follow-up) Follow-up: 17-19 weeks	100 per 1000 Moderate 100 per 1000	33 per 1000 (9 to 119) 33 per 1000 (9 to 119)	RR 0.33 (0.09 to 1.19)	180 (1 study)	⊕⊕⊖⊖ low1,2,

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total population size is less than 400 (a threshold rule-of-thumb)
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Adverse event outcomes (by intervention)

Norethisterone versus placebo

There was low quality evidence for a moderate to large effect on days of vaginal bleeding in favour of placebo compared with norethisterone at the end of intervention (p<0.0001) and short-term follow-up (p<0.0001), and for troublesome bleeding at the end of intervention (p=0.002), however the imprecision of these effect estimates was serious due to the small population and number of events (Table 280). There was no statistically or clinically significant effect of norethisterone on return of sexual interest.

Table 280: Summary of findings table for effects of omega-3 compared with placebo on adverse events

Adverse events: Norethisterone versus placebo for prevention (risk factors present)

Patient or population: patients with prevention (risk factors present)
Settings:
Intervention: Adverse events: Norethisterone versus placebo

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Adverse events: Norethisterone versus placebo				
Vaginal bleeding days (Post-treatment) Follow-up: 6 weeks		The mean vaginal bleeding days (post-treatment) in the intervention groups was 0.74 standard deviations higher (0.43 to 1.06 higher)		164 (1 study)	⊕⊕⊖⊖ low1	SMD 0.74 (0.43 to 1.06)
Vaginal bleeding days (Short-term follow-up, 9-16 weeks post-intervention) Follow-up: 12 weeks		The mean vaginal bleeding days (short-term follow-up, 9-16 weeks post-intervention) in the intervention groups was 0.77 standard deviations higher (0.45 to 1.09 higher)		164 (1 study)	⊕⊕⊖⊖ low1	SMD 0.77 (0.45 to 1.09)
Troublesome bleeding (Post-treatment) Follow-up: 6 weeks	Study population		RR 3.18 (1.53 to 6.57)	165 (1 study)	⊕⊕⊖⊖ low1	
	100 per 1000	318 per 1000 (153 to 657)				
	Moderate					
	100 per 1000	318 per 1000 (153 to 657)				
No return of sexual interest (Post-treatment) Follow-up: 12 weeks	Study population		RR 1.14 (0.88 to 1.46)	149 (1 study)	⊕⊕⊖⊖ low1	
	583 per 1000	665 per 1000 (513 to 852)				
	Moderate					
	583 per 1000	665 per 1000 (513 to 851)				

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total population size is less than 400 (a threshold rule-of-thumb)

8.2.5 Clinical evidence for preventative effects on outcomes - prophylaxis of mental health problems

Summary of findings can be found in the tables presented in this section. The full GRADE evidence profiles and associated forest plots can be found in Appendix 22 and Appendix 19, respectively.

Recurrence of depression outcomes (by intervention)

SSRIs (sertraline) versus placebo

There was low quality, single study (N=22) evidence for a large beneficial effect of SSRIs on preventing recurrence of depression at post-treatment (p=0.06). However, the imprecision of this effect estimate was very serious due to the very small population size and large 95% CIs (Table 281).

Table 281: Summary of findings table for effects of SSRIs (sertraline) compared with placebo on depression outcomes

Depression recurrence: SSRI (Sertraline) versus placebo for prophylaxis of mental health disorders

Patient or population: patients with prophylaxis of mental health disorders
Settings:
Intervention: Depression: SSRI (Sertraline) versus placebo

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Depression recurrence: SSRI (Sertraline) versus placebo				
Recurrence of depression (post-treatment)	Study population		**RR 0.14** (0.02 to 1.07)	22 (1 study)	⊕⊖⊖⊖ **very low**[1,2]	
	500 per 1000	70 per 1000 (10 to 535)				

HRSD >=15 on two occasions and DSM-IV Follow-up: 17 weeks	Moderate	
	500 per 1000	70 per 1000 (10 to 535)

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Unclear attrition bias and independence of data assumption contravened
2 Total population size is less than 400 (a threshold rule-of-thumb) and 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

TCAs (nortriptyline) versus placebo

There was no evidence for a statistically or clinically significant benefit of nortriptyline on recurrence of depression at post-treatment (p=0.94) or long-term follow-up (p=0.63,
Table 282)

Table 282: Summary of findings table for effects of TCAs (nortriptyline) compared with placebo on depression outcomes

Depression: TCA versus placebo for prophylaxis of mental health disorders

Patient or population: patients with prophylaxis of mental health disorders
Settings:
Intervention: Depression: TCA versus placebo

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk Control	Corresponding risk Depression: TCA versus placebo				
Recurrence of major depression (post-treatment) HRSD >=15 and RDC for major depression Follow-up: 22 weeks	Study population		**RR 0.96** (0.36 to 2.59)	51 (1 study)	⊕⊕⊖⊖ **low**[1]	
	240 per 1000	230 per 1000 (86 to 622)				
	Moderate					
	240 per 1000	230 per 1000 (86 to 622)				

Recurrence of major depression postpartum (long-term follow-up, 25-103 weeks post-intervention) HRSD >=15 and RDC for major depression Follow-up: 26 weeks	Study population		**RR 1.2** (0.57 to 2.55)	51 (1 study)	⊕⊕⊖⊖ **low**[1]
	320 per 1000	384 per 1000 (182 to 816)			
	Moderate				
	320 per 1000	384 per 1000 (182 to 816)			

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total population size is less than 400 (a threshold rule-of-thumb) and 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Adverse events outcomes (by intervention)

SSRIs (sertraline) versus placebo

There was very low quality, single study (N=22) evidence for a statistically significant increased risk drowsiness with SSRIs (sertraline, p=0.002), however the imprecision of this effect estimate was very serious due to the very small population size and large 95% CI (Table 283). There was no evidence for an effect of SSRIs (sertraline) on dizziness.

Table 283: Summary of findings table for effects of SSRIs (sertraline) compared with placebo on adverse events

Adverse events: SSRI (Sertraline) versus placebo for prophylaxis of mental health disorders						
Patient or population: patients with prophylaxis of mental health disorders Settings: Intervention: Adverse events: SSRI (Sertraline) versus placebo						
Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Adverse events: SSRI (Sertraline) versus placebo				
Dizziness (post-treatment) Follow-up: 17 weeks	Study population		RR 4.57 (0.69 to 30.22)	22 (1 study)	⊕⊖⊖⊖ very low1,2	
	125 per 1000	571 per 1000 (86 to 1000)				
	Moderate					

	125 per 1000	571 per 1000 (86 to 1000)			
Drowsiness (post-treatment) Follow-up: 17 weeks	Study population		RR 1.93 (1 to 3.74)	22 (1 study)	⨁⊖⊖⊖ very low1,2
	500 per 1000	965 per 1000 (500 to 1000)			
	Moderate				
	500 per 1000	965 per 1000 (500 to 1000)			

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Unclear attrition bias and independence of data assumption contravened
2 Total population size is less than 400 (a threshold rule-of-thumb) and 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

TCAs (nortriptyline) versus placebo

There was low quality, single study evidence (N=51) for a large effect of nortriptyline on the number of participants reporting constipation at post-treatment ($p<0.01$), however imprecision of this effect estimate was very serious due to the small population size and large 95% CI (Table 284). There was no statistically or clinically significant effect of nortriptyline on discontinuation due to adverse effects at post-treatment ($p=0.48$).

Table 284: Summary of findings table for effects of TCAs (nortriptyline) compared with placebo on adverse events

Adverse events: TCA (Nortriptyline) versus placebo for prophylaxis of mental health disorders

Patient or population: patients with prophylaxis of mental health disorders
Settings:
Intervention: Adverse events: TCA (Nortriptyline) versus placebo

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Adverse events: TCA (Nortriptyline) versus placebo				
Discontinuation due to adverse events (post-treatment) Follow-up: 20 weeks	Study population		RR 0.32 (0.01 to 7.53)	51 (1 study)	⊕⊕⊖⊖ low1	
	40 per 1000	13 per 1000 (0 to 301)				
	Moderate					
	40 per 1000	13 per 1000 (0 to 301)				
Constipation (post-treatment) Follow-up: 20 weeks	Study population		RR 3.21 (1.55 to 6.64)	51 (1 study)	⊕⊕⊖⊖ low1	
	240 per 1000	770 per 1000 (372 to 1000)				
	Moderate					

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total population size is less than 400 (a threshold rule-of-thumb) and 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Leaving the study early (by intervention)

SSRIs (sertraline) versus placebo

There was no statistically or clinically significant difference between of SSRIs (sertraline) and placebo on leaving the study early for any reason except for recurrence at post-treatment (p=0.17, Table 285).

Table 285: Summary of findings table for effects of SSRIs (sertraline) compared with placebo on leaving the study early

Leaving the study early: SSRI (Sertraline versus placebo) for prophylaxis of mental health disorders

Patient or population: patients with prophylaxis of mental health disorders
Settings:
Intervention: Leaving the study early: SSRI (Sertraline versus placebo)

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Leaving the study early: SSRI (Sertraline versus placebo)				
Leaving study early for any reason except recurrence (post-treatment) Follow-up: 17 weeks	Study population		RR 3.76 (0.56 to 25.21)	25 (1 study)	⊕⊖⊖⊖ very low1,2	
	125 per 1000	470 per 1000 (70 to 1000)				
	Moderate					
	125 per 1000	470 per 1000 (70 to 1000)				

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Unclear attrition bias and independence of data assumption contravened
2 Total population size is less than 400 (a threshold rule-of-thumb) and 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

TCAs (nortriptyline) versus placebo

There was statistically or clinically significant difference between TCAs (nortriptyline) and placebo on leaving the study early for any reason except for recurrence at post-treatment (p=0.63,)

Table 286: Summary of findings table for effects of TCAs (nortriptyline) compared with placebo on adverse events

Leaving the study early: TCA versus placebo for prophylaxis of mental health disorders						
Patient or population: patients with prophylaxis of mental health disorders **Settings:** **Intervention: Leaving the study early: TCA versus placebo**						
Outcomes	**Illustrative comparative risks* (95% CI)** **Assumed risk** **Control**	**Corresponding risk** **Leaving the study early: TCA versus placebo**	**Relative effect (95% CI)**	**No of Participants (studies)**	**Quality of the evidence (GRADE)**	**Comments**
Leaving study early for any reason except recurrence Follow-up: 20 weeks	Study population 185 per 1000	137 per 1000 (41 to 461) Moderate	RR 0.74 (0.22 to 2.49)	56 (1 study)	⊕⊕⊖⊖ low1	
	185 per 1000	137 per 1000 (41 to 461)				
*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).						
GRADE Working Group grades of evidence **High quality**: Further research is very unlikely to change our confidence in the estimate of effect. **Moderate quality**: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate. **Low quality**: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate. **Very low quality**: We are very uncertain about the estimate.						
1 Total population size is less than 400 (a threshold rule-of-thumb) and 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)						

8.2.6 Health economic evidence

Systematic literature review

No studies assessing the cost effectiveness of pharmacological interventions for the prevention of mental health problems in pregnancy or the postnatal period were

identified by the systematic search of the economic literature undertaken for this guideline. Details on the methods used for the systematic search of the economic literature are described in Chapter 3.

8.3 PHARMACOLOGICAL INTERVENTIONS FOR THE TREATMENT OF MENTAL HEALTH PROBLEMS IN PREGNANCY AND THE POSTNATAL PERIOD

8.3.1 Clinical review protocol (treatment)

The review protocol summary, including the review question(s) and the eligibility criteria used for this section of the guideline, can be found in Table 287.A complete list of review questions can be found in Appendix 8; further information about the search strategy can be found in Appendix 10; the full review protocols can be found in Appendix 9.

The review strategy was to evaluate the clinical effectiveness of the interventions using meta-analysis. However, in the absence of adequate data, the available evidence was synthesised using narrative methods. An analysis of all interventions was conducted and graded. Where possible both an available case analysis and an intention-to-treat (ITT) analysis (last observation carried forward [LOCF]; worst case scenario [WCS]) were used.

Table 287: Clinical review protocol summary for the review of pharmacological interventions for the treatment of mental health problems

Component	Description
Review question(s)	RQ 4.2 For women with mental health problems who are pregnant or postnatal, what are the benefits and/or potential harms of pharmacological interventions to treat mental health problems? RQ 4.3 For women with mental health problems who are pregnant or postnatal, what are the benefits and/or potential harms of combined pharmacological and psychosocial treatment interventions to treat mental health problems?
Population	Included Women who have mental health problems during pregnancy and postnatal period (from delivery to the end of the first year). Include:-Women with sub-threshold symptoms (but no formal diagnosis of a mental health problem) Women with a formal diagnosis of mild, moderate and severe disorders Exclude women:- who are not pregnant or postnatal period (up to 1 year postnatal)
Intervention(s)	Pharmacological interventions, including: Psychotrophic medication Dietary supplements Hormones
Comparison	Any other comparison group, including: Placebo

	Another active intervention
Critical outcomes	Maternal Outcomes Symptom-based Diagnosis of mental health problems Symptomatology Relapse Use of drugs/alcohol Service utilisation Hospitalisation Retention in services (assessed through dropout rates as a proxy measure) Health service utilisation (for instance, use of psychiatric services) Experience of care Satisfaction (validated measures only, specific items will not be analysed) Acceptability of treatment (assessed through questioning or through including dropout as a proxy measure) Quality of life Quality of life measures Functional disability Social functioning Social support Self-esteem Perceived parenting stress Maternal confidence Preservation of rights Harm Side effects (including dropout because of side effects) Maternal mortality and serious morbidity including self-harm and suicide attempts Quality of mother-infant interaction Quality of mother-infant interaction (including maternal sensitivity and child responsivity) Maternal attitude towards motherhood Establishing or continuing breastfeeding Infant outcomes (no restriction on length of follow-up) Fetal and infant physical development (including congenital malformations) Side effects (especially of pharmacological interventions for the fetus and for the infant if breastfeeding) Apgar score Birth weight Admission to neonatal intensive care unit Cognitive development of the infant Emotional development of the infant Physical development of the infant Prevention of neglect or abuse of the infant Optimal care of infant (for example vaccinations, well-baby check-ups) Foetal/infant mortality Foetal/infant morbidity Service use Planned (health visitor, vaccinations, well-baby check-ups) Unplanned (emergency department visits, inpatient, urgent or acute care)

	Social service involvement Mode of delivery
Study design	Systematic reviews of RCTs Primary RCTs For harm review, cohort studies were included
Note. None	

8.3.2 Studies considered (treatment)[14]

Eleven RCTs met the eligibility criteria for this review: APPLEBY1997 (Appleby et al, 1997), BLOCH2012 (Bloch et al., 2012), FREEMAN2008 (Freeman et al., 2008), GREGOIRE1996 (Gregoire et al., 1996), HANTSOO2014 (Hantsoo et al., 2014), MOZURKEWICH2013 (Mozurkewich et al., 2013), REES2008 (Rees et al., 2008), SHARP2010, SU2008 (Su et al., 2008), WISNER2006 (Wisner et al., 2006), YONKERS2008. All of these studies were published in peer-reviewed journals between 1997 and 2014. In addition 11 studies were excluded from the review. The reasons for exclusion were that the studies were not RCTs, insufficient data were provided for extraction and studies were open label. Further information about both included and excluded studies can be found in Appendix 18.

There were four studies that involved a comparison between omega-3 and placebo. Two studies compared SSRIs (one sertraline, and one paroxetine) with placebo, and one study compared SSRIs (sertraline) with TCAs (nortriptyline). One study compared antidepressants (primarily SSRIs) with general standard care. There were two studies that involved a comparison of SSRIs (one fluoxetine, one sertraline) in combination with a psychological intervention (one counselling, one brief dynamic psychotherapy) and one study compared hormones (oestradiol patches) with placebo (Table 288).

For the review of pharmacological treatment for alcohol or substance misuse, one Cochrane review met the eligibility criteria for this review: MINOZZI2008/2013 (Minozzi et al., 2008; Minozzi et al., 2013) (Table 289). An additional Cochrane review was identified by the search, however, no suitable trials were identified by this review and as a result there was no data that could be extracted (SMITH2009 [Smith et al., 2009]). One further systematic review was identified by the search for this review but was excluded as no new data could be extracted (Jones et al., 2012a).

[14] Here and elsewhere in the guideline, each study considered for review is referred to by a study ID in capital letters (primary author and date of study publication, except where a study is in press or only submitted for publication, then a date is not used).

Table 288: Study information table for trials included in the meta-analyses for any pharmacological interventions versus any alternative comparison

	Omega-3 versus placebo	SSRIs versus placebo	SSRIs versus TCA	SSRIs verus general supportive care	SSRIs/ psychological versus placebo/ psychological	Hormones versus placebo
Total no. of trials (k); participants (N)	4 (251)	2 (108)	1 (109)	1 (254)	2 (129)	1 (64)
Study ID	(1) FREEMAN2008 (2) MOZURKEWICH2013 (3) REES2008 (4) SU2008	(1) HANTSOO2014 (2) YONKERS2008	WISNER2006	SHARP2010	(1) APPLEBY1997 (2) BLOCH2012	GREGOIRE1996
Country	(1) US (2) US (3) Australia (4) Taiwan	(1) US (2) US	US	UK	(1) UK (2) Israel	UK
Mean Age of Paricipants (years)	(1) 30 (2) 30 (3) 33 (4) 31	(1) 31 (2) 26	NR	29	(1) 25 (2) NR	31
Timing of intervention[1]	(1) Pregnancy and postnatal (2) Pregnancy (3)-(4) Pregnancy and postnatal	(1)–(2) Postnatal	Postnatal	Postnatal	(1)–(2) Postnatal	Postnatal
Length of intervention (weeks)	(1) 8 (2) 36 (3) 6 (4) 8	(1) 6 (2) 8	8	4[3]	(1) 12 (2) 8	26
Time points[2]	(1) Post-treatment (2) Post-treatment (3) Post-treatment	(1)-(2) Post-treatment	Post-treatment; Intermediate follow-up	Post-treatment	(1) Post-treatment (2) Post-treatment	Post treatment

	(4) Post-treatment					
Setting	(1)-(4) Clinic (primary)	(1)-(2) Clinic (primary)	Clinic (primary)	Clinic (primary)	(1)-(2) Clinic (primary)	Clinic (primary)
Dose	(1) 1.1g of EPA and 0.8g of DHA in a total of 4 capsules a day (2) 900 mg DHA plus 180 mg EPA[2] (3) 6g a day fish oil every two weeks (4) Total daily dosage of omega-3 fatty acid with 2.2g of EPA and 1.2g of DHA	(1) 50 mg (esculating dose) (2) 10 mg (esculating dose)	25 mg/d of SERT or 10 mg/d of NTP	NR	(1) NR (2) 25 mg for 1 week, followed by 50 mg for 3 more weeks	200µg
Intervention	(1)–(4) Omega-3	(1) Sertraline (2) Paroxetine	Sertraline	Antidepressants	(1) Fluoxetine + counselling (2) Sertraline + brief dynamic psychotherapy	Oestradiol patches
Comparison	(1)–(4) Placebo	(1)- (2) Placebo	Nortriptyline	Listening visits	(1) Placebo + counselling (2) Placebo + brief dynamic psychotherapy	Unmarked placebo patches

Note.
[1] Time points: Post-treatment or first measurement; Short-term follow-up (9-16 weeks post-intervention); Intermediate follow-up (17-24 weeks post-intervention); Long-term follow-up (25-103 weeks post-intervention follow-up); Very long-term follow-up (=>104 weeks).
[2] MOZURKEWICH2013 reported data for EPA and DHA compared with placebo seperately. Data have been combined for this analysis as not a relevant distinction to this review.
[3] Only 4 week data used as this was from RCT design

Table 289: Study information table for the systematic review included in the review of pharmacological interventions for substance misuse

Cochrane review	Primary objective	Inclusion criteria	Included studies	Additional studies
MINOZZI2008/2013	Determine the effectiveness of any maintenance treatment alone or in combination with psychosocial intervention on child health status, neonatal mortality, retaining pregnant women in treatment and reducing the use of substances.	Pregnant women who are opiate-addicted	Fischer et al. (1999) Fischer et al. (2006) Jones et al., (2005) MOTHER study (Chisolm et al., 2013; Coyle et al., 2012; Gaalema et al., 2012; Holbrook et al., 2012; Jansson et al., 2011; Jones et al., 2008; Jones et al., 2010; Jones et al., 2012b; Unger et al., 2011; Winklbaur-Hausknost et al., 2013)	None

8.3.3 Clinical evidence for the efficacy of pharmacological interventions for mental health problems in pregnancy and the postnatal period

Summary of findings can be found in the tables presented in this section. The full GRADE evidence profiles and associated forest plots can be found in Appendix 22 and Appendix 19, respectively.

Non-response to treatment (by intervention)

Omega-3 versus placebo

There was very low quality, single study (N=36) evidence for moderate beneficial effects of omega-3 on response to treatment from both an available case and an ITT analysis approach at endpoint (Table 290). However these effects did not reach statistical significance (p=0.09-0.11) and there was very serious imprecision due to the small number of participants and 95% CIs including estimates of no effect and clinically meaningful benefit. There was no statistically or clinically significant benefit of omega-3 on non-remission using either an available case or an ITT (WCS) analysis approach.

Table 290: Summary of findings table for treatment effects of omega-3 versus placebo on response outcomes

Response to treatment: Omega-3 versus placebo for [the treatment of mental health problems in pregnancy and postnatal period]						
Patient or population: patients with [the treatment of mental health problems in pregnancy and postnatal period] **Settings:** **Intervention: Response to treatment: Omega-3 versus placebo**						
Outcomes	**Illustrative comparative risks* (95% CI)** **Assumed risk** **Control**	**Corresponding risk** **Response to treatment: Omega-3 versus placebo**	**Relative effect (95% CI)**	**No of Participants (studies)**	**Quality of the evidence (GRADE)**	**Comments**
Non-response to treatment (Post-treatment)- Available case analysis HRSD <50%	Study population		RR 0.53 (0.24 to 1.15)	24 (1 study)	⊕⊖⊖⊖ very low1,2	
	727 per 1000	385 per 1000 (175 to 836)				
	Moderate					
	727 per 1000	385 per 1000 (174 to 836)				

reduction Follow-up: 8 weeks					
Non-response to treatment (Post-treatment)- ITT analysis HRSD <50% reduction Follow-up: 8 weeks	Study population 833 per 1000	558 per 1000 (350 to 883)	RR 0.67 (0.42 to 1.06)	36 (1 study)	⊕⊖⊖⊖ very low1,2
	Moderate 833 per 1000	558 per 1000 (350 to 883)			
Non-remission to treatment (Post-treatment)- Available case analysis HRSD >7 Follow-up: 8 weeks	Study population 818 per 1000	614 per 1000 (368 to 1000)	RR 0.75 (0.45 to 1.26)	24 (1 study)	⊕⊖⊖⊖ very low1,2
	Moderate 818 per 1000	614 per 1000 (368 to 1000)			
Non-remission to treatment (Post-treatment)-ITT analysis HAM-D >7 Follow-up: 8 weeks	Study population 889 per 1000	720 per 1000 (516 to 1000)	RR 0.81 (0.58 to 1.13)	36 (1 study)	⊕⊖⊖⊖ very low1,2
	Moderate 889 per 1000	720 per 1000 (516 to 1000)			

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Risk of bias due to unclear selection bias, detection bias and attrition bias
2 Total population size is less than 400 (a threshold rule-of-thumb) and 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

SSRIs (sertraline/paroxetine) versus placebo

There were mixed results for treatment effects on response outcomes associated with SSRIs (Table 291). Adopting an available case analysis approach, there was very low quality, single study evidence (N=33) for a large benefit of SSRIs (sertraline) on non-response at endpoint (p=0.05), however there was very serious imprecision due to the small number of participants and events. Using an ITT (LOCF) analysis very low quality evidence from two studies (N=106) found no statistically significant effect on non-remission (p=0.28) although the effect just met the threshold for a clinically appreciable benefit. There was low to very low quality evidence for a statistically significant and moderate effect of SSRIs on non-remission at endpoint using both an available case (p=0.05) and an ITT (LOCF, p=0.04) analysis, however the quality of evidence was very low due to serious imprecision and high risk of attrition bias.

Table 291: Summary of findings table for treatment effects of SRRIs compared with placebo on response outcomes

Response to treatment: SSRIs versus placebo for [the treatment of mental health problems in pregnancy and postnatal period]

Patient or population: patients with [the treatment of mental health problems in pregnancy and postnatal period]
Settings:
Intervention: Response to treatment: SSRIs versus placebo

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Response to treatment: SSRIs versus placebo	Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
Non-response to treatment (Post-treatment)- Available case analysis* >10 HRSD >50% decrease, improvement on CGI Follow-up: 6 weeks	Study population 722 per 1000 Moderate 722 per 1000	332 per 1000 (152 to 722) 332 per 1000 (152 to 722)	RR 0.46 (0.21 to 1)	33 (1 study)	⊕⊕⊖⊖ low1	
Non-response to treatment (Post treatment)- ITT analysis** >10 HRSD, >50% decrease, improvement on CGI or CGI-I=1 or 2 Follow-up: 6-8 weeks	Study population 704 per 1000 Moderate 711 per 1000	521 per 1000 (366 to 746) 526 per 1000 (370 to 754)	RR 0.74 (0.52 to 1.06)	106 (2 studies)	⊕⊖⊖⊖ very low2	
Non-remission (Post-treatment)- Available case analysis HRSD >7 Follow-up: 6 weeks	Study population 778 per 1000 Moderate 778 per 1000	397 per 1000 (202 to 778) 397 per 1000 (202 to 778)	RR 0.51 (0.26 to 1)	33 (1 study)	⊕⊕⊖⊖ low1	
Non-remission (Post-treatment)- ITT analysis	Study population 833 per 1000	583 per 1000 (450 to 758)	RR 0.7 (0.54 to 0.91)	106 (2 studies)	⊕⊖⊖⊖ very low1,2	

HRSD >7 or HRSD >8 Follow-up: 6-8 weeks	Moderate	
	823 per 1000	576 per 1000 (444 to 749)

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total population size is less than 400 (a threshold rule-of-thumb)
2 Risk of bias due to high attrition
* Completer: participants with at least 3 post-randomisation assessments (completer)
** Method of ITT unclear

SSRIs in combination with psychological interventions compared with placebo in combination with psychological interventions

There was low quality, single study (N=42) evidence for a moderate effect of SSRIs combined with brief dynamic psychotherapy on response and remission using an ITT (LOCF/WCS) analysis approach (Table 292), however this was not statistically significant (p=0.2-0.22) and the confidence in the estimate was low due to number of events being less than 300 and 95% CI crosses both line of no effect and measure of appreciable benefit or harm.

Table 292: Summary of findings table for effects of SRRIs in combination with psychosocial interventions compared with placebo in combination with psychosocial interventions on response outcomes

Response to treatment: SSRI/ psychosocial interventions versus placebo/ psychosocial interventions for [the treatment of mental health problems in pregnancy and postnatal period]

Patient or population: patients with [the treatment of mental health problems in pregnancy and postnatal period]
Settings:
Intervention: Response to treatment: SSRI/psychosocial interventions versus placebo/ psychosocial interventions

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	 Corresponding risk Response to treatment: SSRI/Pscy versus placebo/Pscy	Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
Non-response to treatment (Post-treatment)- ITT analysis* (MADRS or EPDS >50%) Follow-up: 8 weeks	Study population		RR 0.6 (0.27 to 1.32)	42 (1 study)	⊕⊕⊖⊖ low1	
	500 per 1000	300 per 1000 (135 to 660)				
	Moderate					
	500 per 1000	300 per 1000 (135 to 660)				
Non-remission to treatment (Post-treatment)-ITT analysis* Follow-up: 8 weeks	Study population		RR 0.64 (0.32 to 1.3)	42 (1 study)	⊕⊕⊖⊖ low1	
	545 per 1000	349 per 1000 (175 to 709)				
	Moderate					
	546 per 1000	349 per 1000 (175 to 710)				

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 number of events is less than 300 (a threshold rule-of-thumb) and 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
*Calculated based on LOCF and WCS for those not included in LOCF

SSRIs versus TCAs

There were inconsistent results for response outcomes associated with SSRIs compared with TCAs. There was no evidence of a statistically or clinically significant effect of SSRIs compared with TCAs on non-response or non-remission using an ITT (LOCF) analysis approach at post-treatment (Table 293). At intermediate follow-up there was a large effect in favour of TCAs on response using an available case analysis, however the confidence in this effect estimate is very low due to very serious imprecision (small event rate and the 95% CI included both no effect and appreciable benefit).

Table 293: Summary of findings table for effects of SRRIs compared with TCAs on response outcomes

Response to treatment: SSRI compared with TCA for [the treatment of mental health problems in pregnancy and postnatal period]

Patient or population: patients with [the treatment of mental health problems in pregnancy and postnatal period]
Settings:
Intervention: Response to treatment: SSRI
Comparison: TCA

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk TCA	Corresponding risk Response to treatment: SSRI				
Non-response to treatment (Post-treatment)-ITT analysis HRDS<50% reduction Follow-up: 8 weeks	Study population		RR 1.39 (0.84 to 2.27)	109 (1 study)	⊕⊖⊖⊖ very low1,2	
	315 per 1000	438 per 1000 (264 to 715)				
	Moderate					
	315 per 1000	438 per 1000 (265 to 715)				
Non-remission to treatment (Post-treatment)-ITT analysis HRDS >7 Follow-up: 8 weeks	Study population		RR 1.05 (0.74 to 1.5)	109 (1 study)	⊕⊖⊖⊖ very low1,2	
	519 per 1000	544 per 1000 (384 to 778)				
	Moderate					
	519 per 1000	545 per 1000 (384 to 779)				
Non-response to treatment (Intermediate follow-up, 17-24 weeks post-intervention)-	Study population		RR 2.81 (0.12 to 63.83)	29 (1 study)	⊕⊖⊖⊖ very low1,2	
	0 per 1000	0 per 1000 (0 to 0)				
	Moderate					

Available case analysis HRDS<50% reduction Follow-up: 22 weeks	0 per 1000	0 per 1000 (0 to 0)			
Non-remission to treatment (Intermediate follow-up, 17-24 weeks post-intervention)- Available case analysis HRDS >7 Follow-up: 22 weeks	Study population		RR 1.24 (0.34 to 4.6)	29 (1 study)	⨁⊖⊖⊖ very low1,2
	214 per 1000	266 per 1000 (73 to 986)			
	Moderate				
	214 per 1000	265 per 1000 (73 to 984)			

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Risk of bias due to incomplete outcome data (discontinuation between groups unbalanced)
2 Total population size is less than 400 (a threshold rule-of-thumb) and 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Depression outcomes (by intervention)

SSRIs versus placebo

There was very low quality, single study (N=31) evidence for a moderate beneficial effect of SSRIs (paroxetine) on mean depression scores at the end of intervention using an available case analysis (p=0.10, Table 294). However, the quality of this evidence was very low due to very serious imprecision (with small number of participants and 95% CIs including estimates of no effect and clinically meaningful benefit) and a high risk of attrition bias.

Table 294: Summary of findings table for effects of SRRIs compared with placebo on depression outcomes

Depression: SSRIs versus placebo for [the treatment of mental health problems in pregnancy and postnatal period]

Patient or population: patients with [the treatment of mental health problems in pregnancy and postnatal period]
Settings:
Intervention: Depression: SSRIs versus placebo

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Depression: SSRIs versus placebo				
Depression mean scores (Post-treatment)- Available case analysis HRDS Follow-up: 6 weeks		The mean depression mean scores (post-treatment)- available case analysis in the intervention groups was 0.6 standard deviations lower (1.33 lower to 0.12 higher)		31 (1 study)	⊕⊖⊖⊖ very low1,2	SMD -0.6 (-1.33 to 0.12)

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Risk of bias due to high attrition
2 Total population size is less than 400 (a threshold rule-of-thumb) and 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

SSRIs versus TCA

There was no evidence for a statistically significant benefit of SSRIs compared with TCAs on mean depression scores using an available analysis approach at post-treatment or at intermediate follow-up (p=0.6-0.88, Table 295).

Table 295: Summary of findings table for effects of SRRIs compared with TCAs on depression outcomes

Depression: SSRI versus TCA for [the treatment of mental health problems in pregnancy and postnatal period]

Patient or population: patients with [the treatment of mental health problems in pregnancy and postnatal period]
Settings:
Intervention: Depression: SSRI versus TCA

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk Depression: SSRI versus TCA	Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
Depression mean scores (Post-treatment)- Available case analysis HRDS Follow-up: 8 weeks		The mean depression mean scores (post-treatment)- available case analysis in the intervention groups was 0.03 standard deviations higher (0.4 lower to 0.47 higher)		83 (1 study)	⊕⊕⊖⊖ low1,2	SMD 0.03 (-0.4 to 0.47)
Depression mean scores (Intermediate follow-up, 17- 24 weeks post intervention)- Available case analysis HRDS Follow-up: 22 weeks		The mean depression mean scores (intermediate follow-up, 17- 24 weeks post intervention)- available case analysis in the intervention groups was 0.2 standard deviations higher (0.53 lower to 0.93 higher)		29 (1 study)	⊕⊖⊖⊖ very low1,3	SMD 0.2 (-0.53 to 0.93)

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Risk of bias due to incomplete outcome data (discontinuation between groups unbalanced)
2 Total population size is less than 400 (a threshold rule-of-thumb)
3 Total population size is less than 400 (a threshold rule-of-thumb) and 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

SSRIs in combination with psychological interventions compared with placebo in combination with psychological interventions

There was low quality evidence for a moderate beneficial effect of SSRIs combined with psychosocial interventions on mean depression scores post-intervention using both an available case (p=0.03) and an ITT (LOCF, p=0.02) analysis (Table 296). However the quality of this evidence was low due to serious imprecision (with small number of participants) and high and unbalanced attrition rates.

Table 296: Summary of findings table for effects of SRRIs in combination with psychosocial interventions compared with placebo in combination with psychosocial interventions on depression outcomes

Depression: SSRI/ psychosocial interventions versus placebo/ psychosocial interventions for [the treatment of mental health problems in pregnancy and postnatal period]

Patient or population: patients with [the treatment of mental health problems in pregnancy and postnatal period]
Settings:
Intervention: Depression: SSRI/ psychosocial interventions versus placebo/ psychosocial interventions

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Depression: SSRI/Psyc versus placebo/Pscy				
Depression mean scores (Post-treatment)- Available case analysis EPDS Follow-up: 12 weeks		The mean depression mean scores (post-treatment)- available case analysis in the intervention groups was 0.56 standard deviations lower (1.07 to 0.04 lower)		61 (1 study)	⊕⊕⊖⊖ low1,2	SMD -0.56 (-1.07 to -0.04)
Depression mean scores (Post-treatment)- ITT analysis EPDS		The mean depression mean scores (post-treatment)- ITT analysis in the intervention groups was 0.42 standard deviations		127 (2 studies)	⊕⊕⊖⊖ low1,2	SMD -0.42 (-0.77 to -0.07)

Follow-up: 8-12 weeks	lower (0.77 to 0.07 lower)

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Risk of bias due to high and unbalanced attrition rate
2 Total population size is less than 400 (a threshold rule-of-thumb)

Antidepressants versus general supportive care

There was very low quality, single study (N=254) evidence for a moderate beneficial effect of antidepressants on depression symptomology at post-treatment using both an available case (p=0.0001) and an ITT (P= 0.0006) analysis (Table 297). There was also a statistically significant beneficial effect favouring antidepressants on mean depression scores using an available case analysis (p=0.0004). However the quality of evidence was very low due to high risk of bias and serious imprecision.

Table 297: Summary of findings table for effects of antidepressants compared with placebo in combination with general supportive care on depression outcomes

Depression: Antidepressants versus general supportive care for [the treatment of mental health problems in pregnancy and postnatal period]

Patient or population: patients with [the treatment of mental health problems in pregnancy and postnatal period]
Intervention: Depression: Antidepressants versus general supportive care

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Depression: Antidepressants versus general supportive care				
Depression symptomology (Post treatment)- Available case	Study population		RR 0.68 (0.56 to 0.83)	218 (1 study)	⊕⊖⊖⊖ very low1,2	
	804 per 1000	546 per 1000 (450 to 667)				
	Moderate					

analysis EPDS >13 Follow-up: mean 4 weeks					
Depression symptomology (Post treatment)-ITT analysis Follow-up: 4 weeks	Study population		RR 0.76 (0.65 to 0.89)	254 (1 study)	⊕⊖⊖⊖ very low1,2
	824 per 1000	626 per 1000 (536 to 733)			
	Moderate				
Depression mean scores (Post-treatment)-Available case analysis Follow-up: 4 weeks		The mean depression mean scores (post-treatment)-available case analysis in the intervention groups was 0.48 standard deviations lower (0.75 to 0.21 lower)		218 (1 study)	⊕⊖⊖⊖ very low1,2

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 High risk of performance bias and only 56% reported taking antidepressants in intervention group
2 Total population size is less than 400 (a threshold rule-of-thumb)

Omega-3 versus placebo

There was no evidence for a statistically or clinically significant effect of omega-3 on mean depression scores using an ITT analysis approach at the end of intervention (Table 298), however there was substantial heterogeneity between the effect sizes of the four studies.

Table 298: Summary of findings table for effects of omega-3 compared with placebo on depression outcomes

Depression: Omega-3/ psychosocial interventions versus placebo/ psychosocial interventions for [the treatment of mental health problems in pregnancy and postnatal period]

Patient or population: patients with [the treatment of mental health problems in pregnancy and postnatal period]
Settings:
Intervention: Depression: Omega-3/ psychosocial interventions versus placebo/ psychosocial interventions

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Depression: Omega-3/psyc versus placebo/pscy				
Depression mean scores (Post-treatment) -ITT analysis EPDS or BDI Follow-up: 6-36 weeks		The mean depression mean scores (post-treatment) -ITT analysis in the intervention groups was 0.08 standard deviations lower (0.61 lower to 0.46 higher)		228 (4 studies)	⊕⊖⊖⊖ very low1,2,3	SMD -0.08 (-0.61 to 0.46)

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Risk of bias due to high attrition and unclear selection bias throughout studies
2 There was evidence of substantial heterogeneity between effect sizes
3 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Hormones (transdermal oestrogen) versus placebo

There was moderate quality, single study (N=64) evidence for a large beneficial effect of hormones (transdermal oestrogen) on mean depression scores using an available case analysis ($p<0.001$) and on symptomology using an ITT analysis

(p<0.0007) at the end of intervention (Table 299). However there was serious imprecision due to the small number of participants and events.

Table 299: Summary of findings table for treatment effects of hormones compared with placebo on depression outcomes

Depression: Hormones versus placebo for [the treatment of mental health problems in pregnancy and postnatal period]

Patient or population: patients with [the treatment of mental health problems in pregnancy and postnatal period]
Settings:
Intervention: Depression: Hormones versus placebo

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Depression: Hormones versus placebo				
Depression symptomology (Post-treatment)- ITT analysis EPDS >=14 Follow-up: 13 weeks	Study population		RR 0.47 (0.3 to 0.74)	64 (1 study)	⊕⊕⊕⊖ moderate1	
	821 per 1000	386 per 1000 (246 to 608)				
	Moderate					
	821 per 1000	386 per 1000 (246 to 608)				
Depression mean scores* (Post-treatment)- Available case analysis EPDS Follow-up: 13 weeks		The mean depression mean scores (post-treatment)- available case analysis in the intervention groups was 1.12 standard deviations lower (1.77 to 0.47 lower)		45 (1 study)	⊕⊕⊕⊖ moderate1	SMD -1.12 (-1.77 to -0.47)

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total population size is less than 400 (a threshold rule-of-thumb)

* No means/SDs given in text, therefore mean EPDS data taken from figure. SDs calculated from SEs taken from same figure (to 1 decimal place).

General mental health outcomes (by intervention)

SSRIs versus TCAs

There was low quality, single study (N=29) for a moderate effect in favour of SSRIs on global severity and improvement symptomology at endpoint using an available case analysis (Table 300). However this effect estimate is low due to very serious imprecision (very small event rate and the 95% CI included both no effect and appreciable benefit, p=0.72). There was no statistically or clinically significant evidence in any effect of SSRIs compared with TCAs on all other general mental health outcomes using an available case analysis at the end of intervention or at intermediate follow-up (p=0.69-0.93,).

Table 300: Summary of findings table for effects SSRIs compared with TCAs on general mental health outcomes

General mental health: SSRI versus TCA for [the treatment of mental health problems in pregnancy and postnatal period]

Patient or population: patients with [the treatment of mental health problems in pregnancy and postnatal period]
Settings:
Intervention: General mental health: SSRI versus TCA

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	Corresponding risk General mental health: SSRI versus TCA	Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
Global assessment of functioning mean score (Post treatment)- Available case analysis Global Assessment scale Follow-up: 8 weeks		The mean global assessment of functioning mean score (post treatment)- available case analysis in the intervention groups was 0.06 standard deviations higher (0.38 lower to 0.49 higher)		83 (1 study)	⊕⊕⊖⊖ low1	SMD 0.06 (-0.38 to 0.49)
	Study population					

Social problems (Post-treatment)- Available case analysis Social problems questionnaire Follow-up: 8 weeks	489 per 1000	445 per 1000 (279 to 710)	RR 0.91 (0.57 to 1.45)	83 (1 study)	⊕⊕⊖⊖ low1	
	Moderate					
	489 per 1000	445 per 1000 (279 to 709)				
Global assessment of functioning mean score (Intermediate follow-up, 17-24 weeks)- Available case analysis Global Assessment scale Follow-up: 22 weeks		The mean global assessment of functioning mean score (intermediate follow-up, 17-24 weeks)- available case analysis in the intervention groups was 0.03 standard deviations higher (0.69 lower to 0.76 higher)		29 (1 study)	⊕⊕⊖⊖ low1	SMD 0.03 (-0.69 to 0.76)
Social problems (Intermediate follow-up, 17-24 weeks)- Available case analysis Social problems questionnaire Follow-up: 22 weeks	Study population		RR 0.93 (0.29 to 3.03)	29 (1 study)	⊕⊕⊖⊖ low1	
	286 per 1000	266 per 1000 (83 to 866)				
	Moderate					
	286 per 1000	266 per 1000 (83 to 867)				
Global severity and improvement symptomology (Post-treatment)- Available case analysis CGI >=4 Follow-up: 8 weeks	Study population		RR 0.65 (0.06 to 6.92)	83 (1 study)	⊕⊕⊖⊖ low1	
	43 per 1000	28 per 1000 (3 to 294)				
	Moderate					
	43 per 1000	28 per 1000 (3 to 298)				

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total population size is less than 400 (a threshold rule-of-thumb) and 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

SSRIs combined with psychosocial interventions versus placebo combined with psychosocial interventions

There was moderate quality, single study evidence (N=40) for a large beneficial effect of SSRIs combined with psychosocial interventions on mean global severity scores ($p<0.01$) using an ITT analysis post-intervention (Table 301). However there was no statistically or clinically significant benefit on mean global improvement, mean distress or mean well-being scores post-treatment (p=0.36-0.63).

Table 301: Summary of findings table for effects of SSRIs combined with psychological interventions compared with placebo combined with psychological interventions on general mental health outcomes

General mental health: SSRI/Pscy versus placebo/Psychotherapy for [the treatment of mental health problems in pregnancy and postnatal period]

Patient or population: patients with [the treatment of mental health problems in pregnancy and postnatal period]
Settings:
Intervention: General mental health: SSRI/Pscy versus placebo/Psychotherapy

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Control	 Corresponding risk General mental health: SSRI/Pscy versus placebo/Psyc	Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
Global severity mean scores (Post-treatment)- ITT analysis CGI mean Follow-up: 8 weeks		The mean global severity mean scores (post-treatment)- ITT analysis in the intervention groups was 1.37 standard deviations lower (2.06 to 0.67 lower)		40 (1 study)	⊕⊕⊕⊖ moderate1	SMD -1.37 (-2.06 to -0.67)
Global Improvement mean scores (Post-treatment)- ITT analysis CGI mean Follow-up: 8 weeks		The mean global improvement mean scores (post-treatment)- ITT analysis in the intervention groups was 0.29 standard deviations lower (0.91 lower to 0.33 higher)		40 (1 study)	⊕⊕⊖⊖ low1,2	SMD -0.29 (-0.91 to 0.33)
Distress mean scores (Post-		The mean distress mean scores (post-treatment)-		40 (1 study)	⊕⊕⊖⊖ low2	SMD -0.15 (-0.77 to 0.47)

treatment)- ITT analysis Mental Health Inventory Follow-up: 8 weeks	ITT analysis in the intervention groups was 0.15 standard deviations lower (0.77 lower to 0.47 higher)			
Well being mean scores (Post-treatment)- ITT analysis Mental Health Inventory	The mean well being mean scores (post-treatment)- ITT analysis in the intervention groups was 0.21 standard deviations higher (0.41 lower to 0.83 higher)	40 (1 study)	⊕⊕⊖⊖ low2	SMD 0.21 (-0.41 to 0.83)

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total population size is less than 400 (a threshold rule-of-thumb)
2 Total population size is less than 400 (a threshold rule-of-thumb) and 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

SSRIs compared with placebo

There was low quality, single study evidence (N=31) for a large beneficial effect of SSRIs on mean global severity and improvement scores (p=0.02) using an available case analysis at the end of intervention (Table 302). However the precision was poor and there are risk of bias concerns with this study due to high rate of attrition.

Table 302: Summary of findings table for effects of SSRIs compared with placebo on general mental health outcomes

General mental health: SSRIs versus placebo for [the treatment of mental health problems in pregnancy and postnatal period]

Patient or population: patients with [the treatment of mental health problems in pregnancy and postnatal period]
Settings:
Intervention: General mental health: SSRIs versus placebo

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	General mental health: SSRIs versus placebo				
Global severity and improvement mean scores- (Post treatment)- Available case analysis CGI Follow-up: 8 weeks		The mean global severity and improvement mean scores- (post treatment)- available case analysis in the intervention groups was 0.9 standard deviations lower (1.65 to 0.16 lower)		31 (1 study)	⊕⊕⊖⊖ low1,2	SMD -0.9 (-1.65 to -0.16)

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Risk of bias due to high attrition
2 Total population size is less than 400 (a threshold rule-of-thumb)

Service utilisation outcomes

SSRIs combined with psychosocial interventions compared with placebo combined with psychosocial interventions (by intervention)

There was no evidence for a clinically or statistically significant benefit of SSRIs combined with psychosocial interventions relative to placebo combined with psychosocial interventions on lorazepam use post-treatment (p=0.34; Table 303).

Table 303: Summary of findings table for effects of SSRIs combined with psychological interventions compared with placebo combined with psychological interventions on service utilisation outcomes

Service Utilisation: SSRI/ psychosocial interventions versus placebo/ psychosocial interventions for [the treatment of mental health problems in pregnancy and postnatal period]

Patient or population: patients with [the treatment of mental health problems in pregnancy and postnatal period]
Settings:
Intervention: Service Utilisation: SSRI/ psychosocial interventions versus placebo/ psychosocial interventions

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Service Utilisation: SSRI/Pscy versus placebo/Pscy				
Lorazepam use (Post-treatment)- ITT analysis Follow-up: 8 weeks	Study population		RR 0.77 (0.45 to 1.32)	40 (1 study)	⊕⊕⊕⊖ moderate1,2	
	650 per 1000	500 per 1000 (292 to 858)				
	Moderate					
	650 per 1000	500 per 1000 (292 to 858)				
	98 per 1000	169 per 1000 (59 to 485)				
	Moderate					
	98 per 1000	170 per 1000 (59 to 487)				

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the **estimate of effect and is likely to change the estimate.**
Very low quality: We are very uncertain about the estimate.

1 Total population size is less than 400 (a threshold rule-of-thumb) and 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)
2 Total population size is less than 400 (a threshold rule-of-thumb)

Omega-3 compared with placebo

There was low quality, single study (N=118) evidence for a moderate effect favouring placebo relative to omega-3 on antidepressant use post-treatment (p=0.31; Table 304). However the confidence in this effect is low due to poor precision (small population and number of events and the 95% CI crosses both line of no effect and measure of appreciable benefit or harm).

Table 304: Summary of findings table for effects of omega-3 compared with placebo on service utilisation outcomes

Service Utilisation: Omega-3 versus placebo for [the treatment of mental health problems in pregnancy and postnatal period]

Patient or population: patients with [the treatment of mental health problems in pregnancy and postnatal period]
Settings:
Intervention: Service Utilisation: Omega-3 versus placebo

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Service Utilisation: Omega-3 versus placebo				
Antidepressant use (Post-treatment)- ITT analysis Follow-up: 26-36 weeks	Study population		RR 1.73 (0.6 to 4.97)	118 (1 study)	⊕⊕⊖⊖ low1	
	98 per 1000	169 per 1000 (59 to 485)				
	Moderate					
	98 per 1000	170 per 1000 (59 to 487)				

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total population size is less than 400 (a threshold rule-of-thumb) and 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

SSRIs compared with placebo

There was low quality, single study (N=36) evidence for increased benzodiazepine use associated with SSRIs (sertraline) at the end of intervention (p=0.14; Table 305). However confidence in this effect is low due to very serious imprecision (the population size and number of events was low and the 95% CI crosses both line of no effect and measure of appreciable benefit or harm).

Table 305: Summary of findings table for effects of SSRIs compared with placebo on service utilisation

Service utilisation: SSRIs versus placebo for [the treatment of mental health problems in pregnancy and postnatal period]

Patient or population: patients with [the treatment of mental health problems in pregnancy and postnatal period]
Settings:
Intervention: Service utilisation: SSRIs versus placebo

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk Control	Corresponding risk Service utilisation: SSRIs versus placebo				
Benzodiazepine use (Post-treatment)- ITT analysis - Sertraline versus placebo	Study population		RR 0.42 (0.13 to 1.33)	36 (1 study)	⊕⊕⊖⊖ low1	
	421 per 1000	177 per 1000 (55 to 560)				
	Moderate					
	421 per 1000	177 per 1000 (55 to 560)				

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total population size is less than 400 (a threshold rule-of-thumb) and 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Leaving the study early (by intervention)

SSRIs combined with psychological interventions compared with Placebo combined with psychological interventions

There was low quality, single study (N=128) evidence for a large effect of leaving the study early due to adverse events in favour of SSRIs combined with psychological interventions (Table 306), however the imprecision is very serious due to very small number of events and the 95% CI crosses both line of no effect and measure of appreciable benefit or harm. There was no statistically or clinically significant effect on leaving the study early due to any other reasons.

Table 306: Summary of findings table for effects of SSRIs combined with psychological interventions compared with placebo combined with psychological interventions on leaving the study early

Leaving the study early: SSRI/Psyc compared with placebo/Psychological interventions for [the treatment of mental health problems in pregnancy and postnatal period]

Patient or population: patients with [the treatment of mental health problems in pregnancy and postnatal period]
Settings:
Intervention: Leaving the study early: SSRI/Psychological
Comparison: placebo/Psychological

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Placebo/Pscy	Leaving the study early: SSRI/Psyc				
Leaving the study due to adverse events (Post-treatment)- Available case analysis Follow-up: 12 weeks	Study population		RR 0.33 (0.04 to 3.08)	86 (1 study)	⊕⊕⊖⊖ low1	
	70 per 1000	23 per 1000 (3 to 215)				
	Moderate					
	70 per 1000	23 per 1000 (3 to 216)				
Leaving study early for any reason (Post-treatment)- Available case analysis Follow-up: 8-12 weeks	Study population		RR 1.22 (0.68 to 2.18)	128 (2 studies)	⊕⊕⊖⊖ low1	
	231 per 1000	282 per 1000 (157 to 503)				
	Moderate					
	208 per 1000	254 per 1000 (141 to 453)				

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total population size is less than 400 (a threshold rule-of-thumb) and 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

SSRIs compared with placebo

There was no statistically or clinically significant effect of SSRIs on leaving the study early due to any reason at endpoint (Table 307).

Table 307: Summary of findings table for effects of SSRIs compared with placebo on leaving the study early

Leaving the study early: SSRIs versus placebo for [the treatment of mental health problems in pregnancy and postnatal period]

Patient or population: patients with [the treatment of mental health problems in pregnancy and postnatal period]
Settings:
Intervention: Leaving the study early: SSRIs versus placebo

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Leaving the study early: SSRIs versus placebo				
Leaving the study early for any reason (Post-treatment)- Available case analysis Follow-up: 6-8 weeks	Study population		RR 0.89 (0.54 to 1.33)	106 (2 studies)	⊕⊕⊖⊖ low1	
	444 per 1000	396 per 1000 (240 to 591)				
	Moderate					
	379 per 1000	337 per 1000 (205 to 504)				

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.

Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total population size is less than 400 (a threshold rule-of-thumb) and 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

SSRIs compared with TCAs

There was low quality, single study (N=109) evidence in favour of TCAs for leaving the study early (Table 308, p=0.06). However the quality of evidence is low due to very serious imprecision (small number of events and 95% CI crosses both line of no effect and measure of appreciable benefit or harm).

Table 308: Summary of findings table for effects of SSRIs compared with TCAs on leaving the study early

Leaving the study early: SSRI compared with TCA for [the treatment of mental health problems in pregnancy and postnatal period]

Patient or population: patients with [the treatment of mental health problems in pregnancy and postnatal period]
Settings:
Intervention: Leaving the study early: SSRI
Comparison: TCA

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk TCA	Corresponding risk Leaving the study early: SSRI				
Leaving the study early for any reason (Post-treatment)- Available case analysis	Study population		RR 1.74 (0.99 to 3.06)	109 (1 study)	⊕⊕⊖⊖ low1	
	241 per 1000	419 per 1000 (238 to 737)				
	Moderate					
	241 per 1000	419 per 1000 (239 to 737)				

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total population size is less than 400 (a threshold rule-of-thumb) and 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Hormones versus placebo

There was low quality, single study (N=64) evidence for fewer participants leaving the study early for any reason in favour of oestradiol (Table 309, p=0.14), however the quality of evidence is low due to very serious imprecision (small number of events and 95% CI crosses both line of no effect and measure of appreciable benefit or harm).

Table 309: Summary of findings table for effects of hormones compared with placebo on service utilisation

Leaving the study early: Hormones versus placebo for [the treatment of mental health problems in pregnancy and postnatal period]

Patient or population: patients with [the treatment of mental health problems in pregnancy and postnatal period]
Settings:
Intervention: Leaving the study early: Hormones versus placebo

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Leaving the study early: Hormones versus placebo				
Leaving study early for any reason (Post-treatment)- Available case analysis	Study population		RR 0.57 (0.26 to 1.22)	64 (1 study)	⊕⊕⊖⊖ low1	
	393 per 1000	224 per 1000 (102 to 479)				
	Moderate					
	393 per 1000	224 per 1000 (102 to 479)				

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total population size is less than 400 (a threshold rule-of-thumb) and 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Omega-3 versus placebo

There was low quality evidence from four studies (N=239) for a moderate beneficial effect of omega-3 on leaving the study early (

Table 310, p=0.09). However the quality of evidence is low due to very serious imprecision (small number of events and 95% CI crosses both line of no effect and measure of appreciable benefit or harm).

Table 310: Summary of findings table for effects of SSRIs compared with placebo on service utilisation

Leaving the study early: Omega-3/psychological interventions versus placebo/ psychological interventions for [the treatment of mental health problems in pregnancy and postnatal period]

Patient or population: patients with [the treatment of mental health problems in pregnancy and postnatal period]
Settings:
Intervention: Leaving the study early: Omega-3/psychological versus placebo/pscy

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Leaving the study early: Omega-3/psychological versus placebo/ psychological interventions				
Leaving the study early for any reason (Post-treatment)- Available case analysis	Study population		RR 0.62 (0.35 to 1.09)	239 (4 studies)	⊕⊕⊖⊖ low1	
	230 per 1000	143 per 1000 (80 to 251)				
	Moderate					
	243 per 1000	151 per 1000 (85 to 265)				

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total population size is less than 400 (a threshold rule-of-thumb) and 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Adverse events outcomes (by intervention)

SSRI combined with psychosocial interventions compared with placebo compared with psychosocial interventions

There was no statistically or clinically significant effect of SSRIs combined with psychological interventions on mean side effect scores (Table 311). There were two cases of hypomanic switching in the SSRIs combined with psychological intervention group and none in the placebo combined with psychological intervention group.

Table 311: Summary of findings table for treatment effects of SSRIs combined with psychological interventions compared with placebo combined with psychological interventions on adverse events

Adverse events: SSRI/ psychosocial interventions versus placebo/ psychosocial interventions for [the treatment of mental health problems in pregnancy and postnatal period]

Patient or population: patients with [the treatment of mental health problems in pregnancy and postnatal period]
Settings:
Intervention: Adverse events: SSRI/ psychosocial interventions versus placebo/ psychosocial interventions

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Adverse events: SSRI/ psychosocial interventions versus placebo/ psychosocial interventions				
Side effect mean scores (Post treatment)- ITT analysis		The mean side effect mean scores (post treatment)- ITT analysis in the intervention		40 (1 study)	⊕⊕⊖⊖ low1	SMD -0.08 (-0.7 to 0.54)

Udvalg for Kliniske Undersølger Side Effect Rating Scale Follow-up: 8 weeks	groups was 0.08 standard deviations lower (0.7 lower to 0.54 higher)

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total population size is less than 400 (a threshold rule-of-thumb) and 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

Omega-3 versus placebo

There was no statistically or clinically significant effect of omega-3 on mild or transient side effects post-treatment (Table 312, p=0.64). There was one case of hypomanic side effects in the omega-3 group and one case of suicide in the placebo group.

Table 312: Summary of findings table for effects of SSRIs compared with placebo on service utilisation

Adverse events: Omega-3/ psychosocial interventions l versus placebo/ psychosocial interventions for [the treatment of mental health problems in pregnancy and postnatal period]
Patient or population: patients with [the treatment of mental health problems in pregnancy and postnatal period] **Settings:** **Intervention: Adverse events: Omega-3/Psychological versus placebo/Pscy**

Outcomes	**Illustrative comparative risks* (95% CI)** **Assumed risk** **Control**	**Corresponding risk** **Adverse events: Omega-3/Psychological versus placebo/ Psychological**	**Relative effect (95% CI)**	**No of Participants (studies)**	**Quality of the evidence (GRADE)**	**Comments**
Any mild/transient side effects (Post-treatment)- Available case analysis Follow-up: 6-8 weeks	Study population 246 per 1000 Moderate	282 per 1000 (157 to 506)	RR 1.15 (0.64 to 2.06)	118 (4 studies)	⊕⊕⊖⊖ low1	

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Total population size is less than 400 (a threshold rule-of-thumb) and 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

SSRIs versus placebo

The evidence for effects of SSRIs on adverse events was very low (Table 313) due to very serious imprecision (very small number of events and 95% CI crosses both line of no effect and measure of appreciable benefit or harm), however there were moderate effects for decreased appetite and dizziness associated with SSRIs (p=0.65-0.3), and there was a large effect for dry mouth associated with SSRIs (p=0.14).

Table 313: Summary of findings table for effects of SSRIs compared with placebo on adverse events

Adverse events: SSRIs versus placebo for [the treatment of mental health problems in pregnancy and postnatal period]

Patient or population: patients with [the treatment of mental health problems in pregnancy and postnatal period]
Settings:
Intervention: Adverse events: SSRIs versus placebo

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk Control	Corresponding risk Adverse events: SSRIs versus placebo				
Decreased appetite (Post treatment)- Available case analysis Follow-up: 8 weeks	Study population		RR 1.5 (0.27 to 8.43)	70 (1 study)	⊕⊖⊖⊖ very low1,2	
	57 per 1000	86 per 1000 (15 to 482)				
	Moderate					
	57 per 1000	85 per 1000 (15 to 481)				
Diarrhoea (Post treatment)- Available case analysis Follow-up: 6-8 weeks	Study population		RR 1.02 (0.32 to 3.3)	106 (2 studies)	⊕⊖⊖⊖ very low1,2	
	93 per 1000	94 per 1000 (30 to 306)				
	Moderate					
	84 per 1000	86 per 1000 (27 to 277)				
Dizziness (Post treatment)- Available case analysis Follow-up: 8 weeks	Study population		RR 2 (0.54 to 7.37)	70 (1 study)	⊕⊖⊖⊖ very low1,2	
	86 per 1000	171 per 1000 (46 to 632)				
	Moderate					
	86 per 1000	172 per 1000 (46 to 634)				
Headache (Post treatment)- Available case analysis Follow-up: 6-8 weeks	Study population		RR 0.75 (0.37 to 1.49)	106 (2 studies)	⊕⊖⊖⊖ very low1,2	
	241 per 1000	181 per 1000 (89 to 359)				
	Moderate					
	186 per 1000	140 per 1000 (69 to 277)				
	Study population					

Nausea (Post treatment)- Available case analysis Follow-up: 6-8 weeks	111 per 1000	108 per 1000 (39 to 301)	RR 0.97 (0.35 to 2.71)	106 (2 studies)	⊕⊖⊖⊖ very low1,2
	Moderate				
	86 per 1000	83 per 1000 (30 to 233)			
Somnolence (Post treatment)- Available case analysis Follow-up: 8 weeks	Study population		RR 1 (0.32 to 3.15)	70 (1 study)	⊕⊖⊖⊖ very low1,2
	143 per 1000	143 per 1000 (46 to 450)			
	Moderate				
	143 per 1000	143 per 1000 (46 to 450)			
Dry mouth (Post treatment)- Available case analysis	Study population		RR 9 (0.5 to 161.13)	70 (1 study)	⊕⊖⊖⊖ very low1,2
	0 per 1000	0 per 1000 (0 to 0)			
	Moderate				
	0 per 1000	0 per 1000 (0 to 0)			

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Risk of bias due to high attrition
2 Total population size is less than 400 (a threshold rule-of-thumb) and 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

8.3.4 Clinical evidence for the efficacy of pharmacological and psychosocial interventions for drug and alcohol misuse in pregnancy and the postnatal period

Data from the only included Cochrane review (MINOZZI2008/2013) reports evidence from up to two studies (N=151-175) for a moderate benefit of buprenorphine relative to methadone on use of primary substance (RR 1.81 [0.70, 4.69]; p=0.22) and for serious adverse effects for the mother (RR 1.69 [0.75, 3.83]; p=0.21) and for the child (RR 4.77 [0.59, 38.49]; p=0.14). However, these effect estimates were imprecise (low event rate and 95% CI includes no effect and measure of appreciable benefit). There was also evidence from up to two studies (N=150-175) for statistically significant benefits of buprenorphine relative to methadone for birth weight (mean difference -365.45 [-673.84, -57.07; p=0.02) and on non-serious adverse effects for the mother (RR 1.22 [1.07, 1.38]; p=0.003). Conversely, there was evidence from three studies (N=223) for a clinically but not statistically significant difference in favour of methadone for dropout (RR 0.64 [0.41, 1.01]; p=0.056). No clinically or

statistically significant differences were found between methadone and buprenorphine for Appearance Pulse Grimace response Activity Respiration (APGAR) score (mean difference 0.0 [-0.03, 0.03]; p=1.0), number who needed treatment for neonatal abstinence syndrome (RR 1.22 [0.89, 1.67]; p=0.22), mean duration of neonatal abstinence syndrome treatment (mean difference 0.00 [-0.03, 0.03]; p=1.0), total amount of morphine for neonatal abstinence syndrome (mean difference 5.06 [-3.36, 13.47]; p=0.24), length of hospital stay (mean difference 4.01 [-1.29, 9.30]; p=0.14), or non-serious adverse effects for the child (RR 1.08 [0.74, 1.59]; p=0.69).

MINOZZI2008/2013 also reviewed single study data (N=48) comparing methadone to oral slow-release morphine and found evidence for large and statistically significant benefits of morphine on use of substance (RR 2.40 [1.00, 5.77]; p=0.05) but no clinically or statistically significant differences for birth weight (mean difference 124.00 [-186.94, 434.94]; p=0.43), mean duration of neonatal abstinence syndrome (mean difference -5.00 [-10.97, 0.97]; p=0.10), or for nicotine consumption (mean difference 4.43 [-1.47, 10.33]; p=0.14).

The literature search failed to identify any substantial body of high quality evidence for pharmacological interventions for drug and alcohol detoxification in pregnant women. However, the GDG were mindful of the fact that this is an area of major concern for healthcare professionals and pregnant women because of the known harms to the fetus (for example, fetal alcohol syndrome) and wished to make some recommendations for this population. Therefore, given the limitations of the current evidence base, the GDG decided to consult with acknowledged experts in the field. A half-day meeting with the experts was convened specifically to discuss two issues: (1) the desirability and criteria which may determine whether or not to undertake an alcohol or opioid detoxification in pregnancy, and (2) any specific modifications that may need to be made to the detoxification other than already covered in the existing NICE guidelines on *Drug Misuse: Opioid Detoxification* (NICE, 2007b) and *Alcohol-use Disorders: Diagnosis, Assessment and Management of Harmful Drinking and Alcohol Dependence* (NICE, 2011c). The GDG and experts concluded that detoxification should be offered to women in pregnancy and that it should be done in conjunction with a specialist mental health and substance misuse services, but they also recognised that a number of women would not wish to undertake a detoxification and that these women should be offered interventions to reduce their opioid and alcohol intake in pregnancy.

8.3.5 Clinical evidence for the efficacy of pharmacological and psychosocial interventions for sleep problems and insomnia in pregnancy and the postnatal period

The literature search did not identify any high quality studies assessing the efficacy of pharmacological and psychosocial interventions for sleep problems and insomnia in pregnant women. However, the GDG was mindful that the previous 2007 guideline recommended low-dose chlorpromazine or amitriptyline for women with

'serious and chronic problems', for which the data are limited. The GDG was concerned that low-dose TCAs such as amitriptyline are potentially risky because, if there is depression associated with the insomnia, then there may be a risk of overdose (amitriptyline is very toxic in overdose). The GDG also considered the unpleasant side effects associated with chlorpromazine.

The GDG considered the potential risks associated with low-dose chlorpromazine or amitriptyline, the risks associated with the use of sedating drugs such as zopiclone, and the review of harms associated with both antidepressants and antipsychotics (see Section 8.4), and agreed by consensus that promethazine is a safer option for pregnant women. It was in the list of drugs to be included in the literature search for this guideline, is available over the counter and is prescribed for occasional insomnia.

8.3.6 Health economics evidence

Systematic literature review

No studies assessing the cost effectiveness of pharmacological interventions for the treatment of mental health problems in pregnant and breastfeeding women were identified by the systematic search of the economic literature undertaken for this guideline. Details on the methods used for the systematic search of the economic literature are described in Chapter 3.

8.4 HARMS ASSOCIATED WITH SPECIFIC DRUGS IN PREGNANCY AND THE POSTNATAL PERIOD

8.4.1 Clinical review protocol (harms associated with specific drugs)

The review protocol summary, the review question (RQ 4.2) and the eligibility criteria used for this section of the guideline, can be found in Table 287. A complete list of review questions can be found in Appendix 8; further information about the search strategy can be found in Appendix 10; the full review protocols can be found in Appendix 9.

8.4.2 Methodology

The initial search strategy involved searching for existing systematic reviews of RCTs, cohort and case-control studies. If no reviews were found, or the reviews were out of date, a search for individual studies was conducted. In addition to the initial search, a call for evidence to drug companies for relevant studies or reports that were not yet available in published form was sent.

Inclusion criteria

Study design: cohort and case-control study designs were included in the review. Both 'retrospective' cohort studies, that is those that identify subjects from past records describing the exposure received and follows them from the time of those records, and 'prospective' cohort studies, that is those that recruit participants before

any exposure and follows them into the future, were included.Case-control designs compare subjects who have an outcome, and look back retrospectively to compare how frequently an exposure is present in each group.

Comparison group: only studies with an unexposed comparison group were included in the review.

Outcome data: in order to be included studies needed to report data on at least one of the malformations and at least one of the drugs of interest, and provide sufficient data to calculate an effect size. Congenital malformations are defined as any birth defect. Congenital malformations have traditionally been divided into categories of "major" and "minor". In this review, and in accordance with the majority of the studies included, the term congenital malformations is used to refer to both major and minor malformations combined. Evidence for the subordinate group of major malformations was also analysed separately. There is no standard major-minor definition or distinction. However, commonly used distinguishing features include that a major malformation has an adverse effect on the individual's health, functioning or social acceptability while a minor malformation is generally considered to be of limited social or medical significance. Where studies made this major-minor distinction, data was extracted using the categorisation as defined in the paper. Where categorisation was not applied in the original paper, the malformations were categorised consistently with the other studies in the review.

Review criteria

The following criteria were considered when assessing studies reporting on harms associated with specific drugs in pregnancy:

Study design: results from different study designs are expected to differ systematically, resulting in increased heterogeneity. Therefore, cohort and case-control study designs were not combined in a meta-analysis, but conducted separately for each study design.

Comparison group: a distinction was made between disorder specific comparison groups, that is, studies which used as a comparison group, those who were unexposed to the drug of interest but had the same disorder as the exposed group, and a comparison group that consisted of women from the general population. Each study was used in only one analysis, and the disorder specific comparison group was prioritised where studies reported data for both. Where possible, subgroup analyses were conducted for disorder specific comparison groups.

Reporting on specific drugs: the class of drugs was used as a start point. The GDG decided to look at individual drugs where data existed and where there was reason to suspect that there may be an issue with an individual drug. However caution was taken in singling out individual drugs where there was limited data, in order to avoid making risky interpretations.

Timing of exposure: to maximise available data, results were pooled for studies reporting exposure during any trimester (however the majority of studies reported at least first trimester use). A sensitivity analysis for timing was not possible due to inconsistent reporting on specific trimester exposure.

Type of exposure: data for drugs taken in monotherapy were prioritised because this was most meaningful in terms of attributing the specific drug to harm, rather than the use of the drug in combination. However caution was taken when interpreting the data; for many mental health problems (for example bipolar disorder) taking drugs in combination is the norm.

Outcome reporting: the highest order class of harms was used as the main analysis, for instance, where studies report congenital malformations (all malformations), major malformations and minor malformations, primary review of outcomes would focus on congenital malformations as the superordinate class. Where there was a priori evidence for specific adverse events, these were reported, however it was noted that these were not independent of the main class of harms. For instance, in the case of antidepressants there is a priori evidence for septal defects and this evidence will be reviewed, but the GDG were mindful that the different classes of harms were not necessarily independent from each other so that in this example septal defects are a subgroup of cardiac malformations which form a subgroup for major malformations which form a subgroup for congenital malformations. Only studies where an appropriate definition of the class of harms was provided were included in the meta-analysis. Only outcomes which had more than one study or a substantial sample size (equivalent to the sample sizes in the multiple study meta-analyses) were included in the review.

Type of data: unadjusted, rather than adjusted data was used for the following reasons: there is considerable variability over what each study adjusts for; unadjusted data is most consistently reported and allows the maximisation of available data; the use of unadjusted data allows for absolute rates to be calculated from the raw event rates.

Statistical analysis: for dichotomous outcomes the effect size was reported as an odds ratio. The odds ratiois the measure of association between an exposure and an outcome, where the odds ratio represents the odds that an outcome will occur given a particular exposure, compared to the odds of the outcome occurring in the absence of that exposure (Szumilas, 2010). However the GDG were cautious of over-interpretation of odds ratios when the actual event rate is low. Therefore absolute risk for exposed and unexposed groups and the absolute risk differences were reported. The absolute risk refers to the actual event rate; the number of cases in each group divided by the total number in the group and multiplied by 1000, resulting in a value out of 1000. The absolute risk difference was calculated by working out the difference between the absolute rates in the exposed group compared to the unexposed group. The reason for including both odds ratios and absolute risks is that where the actual event rate is low in the general population, an

effect may be statistically significant, however the absolute risk difference very low, which may lead to an over-interpretation of the effect size. It was not appropriate to calculate absolute risk values for studies using a case-control design because of the inflated prevalence of the cases in the population. Therefore, where possible, odds ratios were interpreted along-side the absolute values, which were used to inform the recommendations. Continuous outcomes were reported as standard mean differences.

8.4.3 Systematic reviews considered [15]

From the initial search 13 systematic reviews were identified, however of these, only six met the inclusion criteria. Only for the antidepressant class of drugs were eligible systematic reviews identified. These were: GRIGORIADIS2013A (Grigoriadis et al., 2013a), GRIGORIADIS2013B (Grigoriadis et al., 2013b), GRIGORIADIS2013C (Grigoriadis et al., 2013c), MYLES2013 (Myles et al., 2013), ROSS2013 (Ross et al., 2013), WURST2010 (Wurst et al., 2010). The systematic reviews were used a source to identify relevant primary studies for antidepressants, however they were updated and adapted in line with our inclusion criteria, and an independent meta-analysis was conducted. The GDG did not feel that the existing systematic reviews for any other classes of drugs were of sufficient quality, therefore a search of the primary literature was conducted.

8.4.4 Studies considered[16]

Antidepressants

From the existing systematic reviews, 30 studies met the eligibility criteria for the review of antidepressants: BOUCHER2008 (Boucher et al., 2008), CALDERON-MARGALIT2009 (Calderon-Margalit et al., 2009), CASPER2003 (Casper et al., 2003), CHAMBERS1996 (Chambers et al., 1996), COSTEI2002 (Costei et al., 2002), DAVIS2007 (Davis et al., 2007) DIAV-CITRIN2008A (Diav-Citrin et al., 2008a), EINARSON2009 (Einarson et al., 2009), FERREIRA2007 (Ferreira et al., 2007), GALBALLY2009 (Galbally et al., 2009), KALLEN2004 (Kallen et al., 2004), KALLEN2007 (Kallen et al., 2007), KIELER2012 (Kieler et al., 2012), KORNUM2010 (Kornum et al., 2010), KULIN1998 (Kulin et al., 1998), LAINE2003 (Laine et al., 2003), LEVINSONCASTIEL2006 (Levinson-Castiel et al., 2006), MALM2011 (Malm et al., 2011), MASCHI2008 (Maschi et al., 2008), OBERLANDER2006 (Oberlander et al., 2006), OBERLANDER2008 (Oberlander et al., 2008), PEDERSEN2009 (Pedersen et al., 2009), RAI2013 (Rai et al., 2013), SIMON2002 (Simon et al., 2002), SIVOJELEZOVA2005 (Sivojelezova et al., 2005), SURI2007 (Suri et al., 2007), WEN2006 (Wen et al., 2006), WICHMAN2009 (Wichman et al., 2009), WISNER2009

[15] Here and elsewhere in the guideline, each study considered for review is referred to by a study ID in capital letters (primary author and date of study publication, except where a study is in press or only submitted for publication, then a date is not used).

[16] Here and elsewhere in the guideline, each study considered for review is referred to by a study ID in capital letters (primary author and date of study publication, except where a study is in press or only submitted for publication, then a date is not used).

(Wisner et al., 2009), WOGELIUS2006 (Wogelius et al., 2006). Six studies were included in the existing systematic reviews, however did not provide the relevant data for the current review, or reported on single study outcomes: ALWAN2007 (Alwan et al., 2007), BAKKER2010 (Bakker et al., 2010), CHAMBERS2006 (Chambers et al., 2006), EINARSON2008 (Einarson et al., 2008), RAMOS2008 (Ramos et al., 2008), WILSON2011 (Wilson et al., 2011). In addition 13 studies were excluded from the analysis as that did not meet the criteria for this review. Further information about both the included and excluded studies can be found in appendix 18 and the full methodological checklists can be found in Appendix 17.

Risk of autism was not included as an adverse event in any of the systematic reviews identified, however the GDG felt this was an important outcome to consider. An additional search was therefore conducted for studies reporting on risk autism associated with antidepressant exposure in pregnancy. One study met the eligibility criteria for this review: ELMARROUN2014 (El Marroun et al., 2014). In addition four studies were excluded because they did not have a disorder specific comparison group. Table 314 provides summary information for studies included in the meta-analysis. Further information about both the included and excluded studies can be found in Appendix 18 and the full methodological checklists can be found in Appendix 17.

Antipsychotics

Of the eligible studies, 12 met the inclusion criteria. Due to the limited available evidence for antipsychotics, all studies were included in the meta-analysis (including those with single study outcomes), to maximise the available data: AUERBACH1992 (Auerbach et al., 1992), BODEN2012A (Boden et al., 2012a) BODEN2012B (Boden et al., 2012b), DIAV-CITRIN2005 (Diav-Citrin et al., 2005), HABERMANN2013 (Habermann et al., 2013), JOHNSON2012 (Johnson et al., 2012), LIN2010 (Lin et al., 2010), MCKENNA2005 (McKenna et al., 2005), NEWHAM2008 (Newham et al., 2008), PENG2013 (Peng et al., 2013) REIS2008 (Reis et al., 2008), SADOWSKI2013 (Sadowski et al., 2013).

Table 315 provides summary information for the studies included in the meta-analyses. In addition five studies were excluded from the review. The main reason for exclusion was that the studies did not have an unexposed control group. Further information about both the included and excluded studies can be found in Appendix 18 and the full methodological checklists can be found in Appendix 17. Two studies provided disaggregated data for first generation and second generation antipsychotics, however the GDG felt that there was generally very little drug specificity, therefore the analyses were conducted for all antipsychotics as a class.

Anticonvulsants

Of the eligible studies, there were 34 which met the inclusion criteria: ADAB2004/VINTEN2005 (Adab et al., 2004; Vinten et al., 2005), ARTAMA2005 (Artama et al., 2005), ARTAMA2013 (Artama et al., 2013), BODEN2012A, BORTHEN2011 (Borthen et al., 2011), BROSH2011 (Brosh et al., 2011), BURJA2006 (Burja et al., 2006), CANGER1999 (Canger et al., 1999), CASSINA2013 (Cassina et a., 2013), CHARLTON2011 (Charlton et al., 2011), CHRISTENSEN2013 (Christensen et al., 2013), CUNNINGTON2011 (Cunnington et al., 2011), DIAV-CITRIN2001 (Diav-Citrin et al., 2001), DIAV-CITRIN2008B (Diav-Citrin et al., 2008b), DOLK2008 (Dolk et al., 2008), ERIKSSON2005 (Eriksson et al., 2005), GAILY2004 (Gaily et al., 2004), HERNANDEZ-DIAZ2012 (Hernandez-Diaz et al., 2012), HOLMES2001 (Holmes et al., 2001), HOLMES2008 (Holmes et al., 2008), HVAS2000 (Hvas et al., 2000) JENTINK2010 (Jentink et al., 2010), KAAJA2003 (Kaaja et al., 2003), KANEKO1999 (Kaneko et al., 1999), KINI2007 (Kini et al., 2007), MOLGAARD-NIELSEN2011 (Molgaard-Nielsen et al., 2011), MORROW2006 (Morrow et al., 2006), ORNOY1996 (Ornoy et al., 1996), RIHTMAN2013 (Rihtman et al., 2013), RODRIGUEZ-PINILLA2000 (Rodriguez-Pinilla et al., 2000), SAMREN1999 (Samren et al., 1999), STEEGERS-THEUNISSEN1994 (Steegers-Theunissen et al., 1994), VAJDA2007 (Vajda et al., 2007), VEIBY2013 (Veiby et al., 2013), WERLER2011 (Werler et al., 2011). Summary information for the studies included in the meta-analysis can be found in

Table 316. In addition 12 studies met the inclusion criteria however could not be included in the meta-analysis as the relevant data could not be extracted, the outcomes could not be combined or the data was not disaggregated for individual drug: ADAB2001 (Adab et al., 2001), ALMGREN2009 (Almgren et al., 2009), BROMLEY2013 (Bromley at al., 2013), CUNNINGTON2011, FONAGER2000 (Fonager et al., 2000), FORSBERG2011 (Forsberg et al., 2011), KAAJA2002 (Kaaja et al., 2002), KULAGA2011 (Kulaga et al., 2011), NULMAN1997 (Nulman et al., 1997), LIN2009 (Lin et al., 2009), RODRIGUEZ-PINILLA2008 (Rodriguez-Pinilla et al., 2008), THOMAS2008 (Thomas et al., 2008), VAJDA2004 (Vajda et al., 2004). In addition 25 studies were excluded from the review. The main reason for exclusion was that the studies did not have an unexposed control group. Data was disaggregated for carbamazepine, lamotrigine and valproate as the magnitude of risks and specific abnormalities varies for each anticonvulsant and have different properties. Further information about both the included and excluded studies can be found in Appendix 18 and the full methodological checklists can be found in Appendix 17.

Lithium

There were six studies which met the inclusion criteria for the review: BODEN2012A, CORREA-VILLASENOR1994 (Correa-Villaseñor et al., 1994), CZEIZEL1990 (Czeizel et al., 1990), JACOBSON1992 (Jacobson et al., 1992), KALLEN1983 (Kallen et al., 1983), REIS2008 (Reis et al., 2008). Summary information for the studies included in the meta-analyses can be found in Table 317. In addition four studies were excluded from the review, the reasons for exclusion were that the studies did not have an unexposed control group and that no cases of lithium exposure were found in case-control studies. Further information about the included and excluded studies can be found in Appendix 18 and the full methodological checklists can be found in Appendix 17.

Benzodiazepines and related drugs[17]

There were 18 studies which met the inclusion criteria, however only nine studies provided sufficient data to be included in the meta-analysis: BAN2014 (Ban et al., 2014), CZEIZEL1987 (Czeizel et al., 1987), LAEGREID1990 (Laegreid et al., 1990), LAEGREID1992 (Laegreid et al., 1992), LEPPEE2010 (Leppee et al., 2010), OBERLANDER2008 (Oberlander et al., 2008), ORNOY1998 (Ornoy et al., 1998), PASTUSZAK1996 (Pastuszak et al., 1996), WIKNER2007 (Wikner et al., 2007). Nine studies met the criteria but were not included in the meta-analysis because no relevant data could be extracted or the study only reported single outcomes: BONNOT2001 (Bonnot et al., 2001), CORREA-VILLASENOR1994 (Correa-Villaseñor et al., 1994), CZEIZEL1999 (Czeizel et al., 1999), CZEIZEL2003 (Czeizel et al.,2003), CZEIZEL2004 (Czeizel et al.,2004), DIAV-CITRIN1999 (Diav-Citrin et al., 1999), EROS2002 (Eros et al., 2002), KJAER2007 (Kjaer et al., 2007), WANG2010 (Wang et al., 2010). A summary of the studies included in the meta-analysis can be found in

[17] Benzodiazepines and related drugs also refer to anxiolytics and hypnotics.

Table 318. One study (BAN2014) was in press at the time of the review. Further information about the included and excluded studies can be found in Appendix 18 and the full methodological checklists can be found in Appendix 17.

Stimulants

Of the eligible studies, only one met the inclusion criteria: POTTEGARD2014 (Pottegard et al., 2014). In addition four studies were excluded from the review as they did not have an unexposed control group. Summary information for this study can be found in Table 319. Further information about the included and excluded studies can be found in Appendix 18 and the full methodological checklists can be found in Appendix 17.

Table 314: Study information table for studiesincluded in the meta-analysis for adverse events associated with antidepressant exposure

Study ID	Total no. of studies (31); participants	Country	Mean age of participants (years)	Diagnosis of participants	Study design	Timing of exposure	Drugs examined
BOUCHER2008	146	Canada	29	NR	Retrospective cohort	Any trimester	Citalopram, paroxtine, sertraline, fluoxetine, fluvoxamine, venlafaxine, amitriptyline, trazodone, mirtazapine
CALDERON-MARGALIT2009	2,631	US	NR	NR	Prospective cohort	Any trimester	SSRIs (any)
CASPER2003	44	US	36	Depression	Prospective and retrospective cohort	Any trimester	Sertraline, fluoxetine, paroxetine, fluvoxamine
CHAMBERS1996	390	Canada	31	Depression (76.9%); anxiety (8.1%), panic disorder (6.4 %), bipolar disorder (5.8%), OCD (4.0%)	Prospective cohort	Any trimester	Fluoxetine
COSTEI2002	109	Canada	33	Depression (565), anxiety (31%), anxiety and depression (13%), panic attacks (9%)	Prospective cohort	3rd trimester	Paroxetine
DAVIS2007	SSRI: 9,836 TCA: 49,836	US	NR	NR	Retrospective cohort	Any trimester	SSRIs (any), TCAs (any)
DIAV-CITRIN2008A	Total: 2,276 Paroxetine: 463 Fluoxetine: 346	Israel, Italy, Germany	32	Depression, anxiety, OCD, manic depressive disorder, schizoaffective	Prospective cohort	1st trimester	Paroxetine, fluoxetine

				disorder and eating disorder			
EINARSON2009	1,856	Canada	NR	NR	Prospective cohort	1st trimester	Bupropion, citalopram, escitalopram, fluvoxamine, nefazodone, paroxetine, mirtazepine, fluoxetine, trazodone, venlafaxine, sertaline
ELMARROUN2014[1]	445	Netherlands	Maternal: 29 Child: 6	Depression	Prospective cohort	1st trimester	SSRIs (any)
FERREIRA2007	166	Canada	31	Major depression (41%), mixed disorders (26%), other anxiety disorders 16%), GAD (14%), and unknown (3%)	Retrospective cohort	3rd trimester	Citalopram, fluoxetine, fluvoxamine, paroxetine, sertraline venlafaxine
GALBALLY2009	50	Australia	32	depression	Prospective cohort	At least 3rd trimester	Sertraline, venlafaxine, fluoxetine, citalopram, fluvoxamine, mianserin, mirtazepine, paroxetine
KALLEN2004	583,793	Sweden	NR	NR	Prospetive cohort	At least 3rd trimester	TCAs, SSRIs, and other antidepressants
KALLEN2007	880,431	Sweden	NR	NR	Retrospective cohort	1st trimester	Paroxetine, fluoxetine, citalopram, sertraline, fluvoxamine, escitalopram
KIELER2012	1,618,255	Denmark, Finland, Iceland, Norway, Sweden	NR	NR	Prospective cohort	Any trimester	Fluoxetine, citalopram, paroxetine, sertraline, fluvoxamine, escitalopram

KORNUM2010	215,774	Denmark	30	NR	Retrospective cohort	Any trimester	paroxetine, fluoxetine, sertraline, citalopram, escitalopram,
KULIN1998	Total: 534 Sertraline: 147 Paroxetine: 97 Fluvoxamine: 26	Canada	31	Depression	Prospective cohort	1st trimester	Sertraline, paroxetine, fluvoxamine, fluoxetine
LAINE2003	40	Finland	33	depression (50%), panic disorder (20%)	Prospective cohort	Any trimester	Citalopram, fluoxetine
LEVINSONCASTIEL2006	120	Israel	32	NR	Prospective cohort	At least 3rd trimester	Paroxetine, fluoxetine, citalopram, venlafaxine, sertraline
MALM2011	635,583	Finland	NR	NR	Retrospective cohort	Any trimester	Citalopram, fluoxetine, paroxetine, sertraline, escitalopram, fluvoxamine
MASCHI2008	1,400	Italy	31	depression (77%), anxiety (25%) and panic attacks (7%)	Prospective cohort	Any trimester	Paroxetine, fluoxetine, amitriptyline
OBERLANDER2006	93,643	Canada	30	Depression	Retrospective cohort	Any trimester	Antidepressant (any)
OBERLANDER2008	109945	Canada	30	NR	Retrospective cohort	1st trimester	Paroxetine, citalopram, fluoxetine, sertraline, fluvoxamine, venlafaxine
PEDERSEN2009	494483	Denmark	NR	Depression	Retrospective cohort	1st trimester	Fluoxetine, citalopram, paroxetine, sertraline
RAI2013	Total: 788 Sertraline: 370 TCA: 418	US	NR	NR	Retrospective cohort	Any trimester	Fluoxetine, fluvoxamine, sertraline, paroxetine, amitriptyline, imipramine, doxepin, nortriptyline, protriptyline, desipramine
SIMON2002	Sertraline/ contol: 370	US	NR	NR	Retrospective cohort	Any trimester	Fluoxetine, fluvoxamine, sertraline, paroxetine, amitriptyline, imipramine,

	TCA/control: 418						doxepin, nortriptyline, protriptyline, desipramine
SIVOJELEZOVA2005	341 (citalopram =108, other SSRIs=115)	Canada	NR	Depression	Prospective cohort	At least first trimester (54% continued through-out pregnan-cy)	Citalopram and other SSRIS
SURI2007	44	US	34	Depression	Prospective cohort	Any trimester	Fluoxetine
WEN2006	4,850	Canada	NR	NR	Retrospective cohort	NR	Citalopram, fluoxetine, fluvoxamine, paroxetine, sertraline
WICHMAN2009	44	US	NR	NR	Retrospective cohort	Any trimester	Citalopram, escitalopram, paroxetine, fluoxetine, sertraline, venlafaxine
WISNER2009	107	US	NR	Depression	Prospective cohort	Any trimester	SSRIs (any)
WOGELIUS2006	4,850	Denmark	NR	NR	Retrospective cohort	Any trimester	SSRIs (any)

Note. Abbreviations: NR=Not reported; SSRI= Selective serotonin reuptake inhibitor; TCA=Tricyclic antidepressants
[1] Identified via additional search of primary studies relating to autism

Table 315. Study information table for studies included in the meta-analysis for adverse events associated with antipsychotic exposure

Study ID	Total no. of studies (10); participants	Country	Mean age of participants (years)	Diagnosis of participants	Study design	Timing of exposure	Drugs examined
AUERBACH1992	58 infants/ 54 mothers	Israel	28	64.29% schizophrenia; 7.14% major depression; 7.14% histrionic personality disorder; 7.14% antisocial personality disorder; 7.14% affective disorder; 7.14% bipolar manic	Prospective cohort	3rd trimester	First-generation antipsychotics
BODEN2012A[1]	667/331376[2]	Sweden	59% 25-34	Bipolar disorder	Prospective cohort	Any trimester	Any antipsychotic
BODEN2012B	358203	Sweden	64% 25-34	90.3% any psychiatric diagnosis; 20.9% schizophrenia; 17.6% other nonaffective psychosis; 11.2% bipolar disorder. Non-exposed group: 8.7% any psychiatric diagnosis; 0.03% schizophrenia; 0.1% other nonaffective psychosis; 0.2% bipolar disorder	Prospective cohort	Any trimester	Any antipsychotic
DIAV-CITRIN2005	846	Israel	Median=31	psychosis (33.5%), schizophrenia (10.7%), depression (9.3%). bipolar disorder (4.2%). schizoaffective disorder (1.4%), anxiety (1.4%). panic attacks (0.9%). hyperemesis gravidarum (0.5%), borderline personality (0.5%), suicide attempt (0.5%), substance abuse (0.5%), and Tourette syndrome (0.5%). 36.1% not specified	Prospective cohort	Any trimester	Any antipsychotic
HABERMANN2013	1967	Germany	32	51.4% psychotic disorders (not otherwise specified); 19.2% schizophrenia; 23.7% depression; 4.9%	Prospective cohort	Any trimester	Any antipsychotic

				bipolar affective disorders; anxiety disorders 7%.			
JOHNSON2012	107	UK	26 weeks (infants) 34 years (mothers)	NR	Prospective cohort	Any trimester	Any antipsychotic
LIN2010	4176	Taiwan	3.5% <20; 15.1% 20–24; 33.3% 25–29; 32.9% 30–34; 15.2% and >34 years)	Schizophrenia	Prospective cohort	Any trimester	Any antipsychotic
MCKENNA2005	302	Canada, Israel, UK	NR	29% depression. 24% schizophrenia. 18% bipolar disorder. 2% schizoaffective. 7% psychotic episode, 5% psychotic depression. 2% OCD, 1% PTSD, 1% schizophreniform disorder	Prospective cohort	Any trimester	Second-generation antipsychotics
NEWHAM2008	108	UK	31	NR	Prospective cohort	Any trimester	Any antipsychotic
PENG2013	152	China	30	DSM-IV diagnosis of schizophrenia	Prospective cohort	Any trimester	Any antipsychotic clozapine (n=33), risperidone (n=16), sulpiride (n=13), olanzapine (n=8), and quetiapine (n=6)
REIS2008	976738	Sweden	NR	NR	Prospective cohort	1st trimester	Any antipsychotic
SADOWSKI2013	266	Canada	NR	36.8% bipolar disorder; 27.1% depression; 9.8% anxiety and depression; 9.8% sleep disorders; 3% schizophrenia; 1.5% schizoaffective disorders	Prospective cohort	Any trimester	Any antipsychotic

Note. Abbreviations: NR=Not reported
[1] BODEN2012A also has data for anticonvulsants and lithium
[2] Number using unexposed general population/disordered comparions

Table 316. Study information table for studies included in the meta-analysis for adverse events associated with anticonvulsant exposure

Study ID	Total no. of studies (35); Participants	Country	Mean age of participants (years)	Diagnosis of participants	Study design	Timing of exposure	Drugs examined
ADAB2004/ VINTEN2005	Unclear	UK	NR	Epilepsy	Retrospective cohort	Any trimester	Carbamazepine, Valproate
ARTAMA2005	2,350	Finland	28	Epilepsy	Prospective Cohort	1st trimester	Carbamazepine, Valproate
ARTAMA2013	4,867	Finland	79% 20-34	Epilepsy	Retrospective Cohort	3rd trimester	Carbamazepine, Lamotrigine, Valproate
BODEN2012A1[1]	709	Sweden	59% 25-34	Bipolar disorder	Prospective Cohort	Any trimester	Carbamazepine, Lamotrigine, Valproate
BORTHEN2011	205	Norway	29	Epilepsy	Retrospective cohort	1st trimester	Carbamazepine, Valproate
BROSH2011	100,736	Israel	29	Epilepsy	Retrospective cohort	1st trimester	Valproate
BURJA2006	69	Slovenia	NR	Epilepsy	Retrospective cohort	Any trimester	Carbamazepine
CANGER1999	452	Israel	25	Epilepsy	Prospective Cohort	1st trimester	Carbamazepine, valproate
CASSINA2013	1,177	Italy	33	57.7% depression, 13.9% anxiety	Prospective Cohort	1st trimester	Carbamazepine, Lamotrigine, Valproate
CHARLTON2011	1,446	UK	30	Epilepsy	Retrospective cohort	1st trimester	Carbamazepine, Lamotrigine, Valproate
CHRISTENSEN 2013	655,615	Denmark	39% 26-30	Epilepsy	Retrospective cohort	Any trimester	Carbamazepine, Lamotrigine, Valproate

DIAV-CITRIN2001	420	Israel	30	epilepsy 80.0%, trigeminal neuralgia or psychiatric disorder (nonepileptic) 12.9%, not specified 7.1%	Prospective cohort	1st trimester	Carbamazepine
DIAV-CITRIN2008B	1,469	Israel	30	81.3% convulsive disorders, 18.7% other indications (psychiatri disorders or migraine)	Prospective cohort	1st trimester	Valporate
DOLK2008	85,563	Mixed	29	Epilepsy (17 out of 495 had no record of maternal epilepsy)	Retrospective Case-control	1st trimester	Lamotrigine
ERIKSSON2005	39	Finland	28	Epilepsy	Retrospective cohort	Any trimester	Carbamazepine, Valproate
GAILY2004/KANTOLA-SORSA2007	144	Finland	Mean age of children=7	Epilepsy	Prospective cohort	Any trimester	Carbamazepine, Valproate
HERNANDEZ-DIAZ2012	3,360	US	30	Epilepsy (92%), mood disorders (6%), migraine (1%), and other conditions	Prospective cohort	Any trimester	Carbamazepine, Lamotrigine, Valproate

HOLMES2001	321	US	NR	Epilepsy	Prospective cohort	Any trimester	Carbamazepine
HOLMES2008	206,908	US	NR	Epilepsy	Prospective cohort	1st trimester	Carbamazepine, Lamotrigine, Valproate
HVAS2000	193	Denmark	NR	Epilepsy	Prospective cohort	1st trimester	Carbamazepine Valproate
JENTINK2010	Unclear	Multiple	NR	NR	Retrospective Case-control	1st trimester	Valproate
KAAJA2002	2,001	Finland	29	NR	Prospective cohort	Any trimester	Carbamazepine
KAAJA2003	790	Finland	29	NR	Prospective cohort	1st trimester	Carbamazepine, Valproate
KANEKO1999	337	Multiple	27	NR	Prospective cohort	1st trimester	Carbamazepine, Valproate
KINI2007	77	UK	NR	Epilepsy	Prospective cohort	Any trimester	Carbamazepine, Valproate
MOLGAARD-NIELSEN2011	837,795	Denmark	45% 25-29	Epilepsy	Prospective cohort	1st trimester	Lamotrigine
MORROW2006	3,607	UK	NR	Epilepsy	Prospective cohort	1st trimester	Carbamazepine, Lamotrigine, Valproate
ORNOY1996	94	Israel	Children 6m-6yrs	Epilepsy	Prospective cohort	Any trimester	Carbamazepine
RIHTMAN2013	124	Iceland	34	NR	Retrospective cohort	1st trimester	Lamotrigine, Valproate
RODRIGUEZ-PINILLA2000	44,241	Spain	NR	NR	Retrospective Case-control	1st trimester	Valporate
SAMREN1999	3,411	Netherlands	41% 25-29	NR	Retrospective cohort	1st trimester	Carbamazepine, Valproate
STEEGERS-THEUNISSEN1994	119	Netherlands	29	Epilepsy	Prospective cohort	Any trimester	Carbamazepine, Valproate

VAJDA2007	546 (234 CBZ; 146 LMG; 166 VPA)	Australia	31	Epilepsy	Prospective cohort	1st trimester	Carbamazepine, Lamotrigine, Valproate
VEIBY2013	726	Norway	NR	Epilepsy	Prospective cohort	Any trimester	Carbamazepine, Lamotrigine, Valproate
WERLER2011	8,554 [26 (CMZ; 14, LMG; 5; VPA; 17)]	US	NR	Epilepsy	Retrospective Case-control	1st trimester	Carbamazepine, Lamotrigine, Valproate

Note. Abbreviations: NR=Not reported; CBZ=carbamazepine; LMG=lamotrigine, VPA=valproate
[1] BODEN2012A also has data for antipsychotics and lithium

Table 317. Study information table for studies included in the meta-analysis for adverse events associated with lithium exposure

Study ID	Total no. of studies (6); participants	Country	Mean age of participants (years)	Diagnosis of participants	Study design	Timing of exposure	Drugs examined
BODEN2012A[1]	661	Sweden	58.5% 25-34	Bipolar disorder	Prospective Cohort	Any trimester	Lithium
CORREA-VILLASENOR1994	6,947	US	31.68% ≥30	NR	Retrospective Case-control	NR	Lithium
CZEIZEL1990	32,244	Hungary	25	NR	Retrospective Case-control	NR	Lithium
JACOBSON1992	186	US	30	NR	Prospective cohort	1st trimester	Lithium
KALLEN1983	121	Sweden	NR	NR	Retrospective cohort	NR	Lithium
REIS2008[3]							Lithium

Note. Abbreviations: NR=Not reported
[1] BODEN2012A also has data for antipsychotics and anticonvulsants

Table 318. Study information for studies included in the meta-analyses for benzodiazepines and related drugs

Study ID	Total no. of studies (8); participants	Country	Mean age of participants (years)	Diagnosis of participants	Study design	Timing of exposure	Drugs examined
BAN2014	21,137	UK	Median: 29	Depression and/or anxiety	Prospective cohort	1st trimester	Diazepam, temazepam and zopiclone
CZEIZEL1987	2,402	Hungary	NR	NR	Retrospective Case-control	Any trimester	Chlordiazepoxide, diazepam and nitrazepam
LAEGREID1990	78	Sweden	NR	NR	Retrospective Case-control	Any trimester	Oxazepam, phenobarbitone, levothyroxine, Nitrofuration, diazepam
LAEGREID1992	46	Sweden	NR	87.5% anxiety disorder; 12.5% depression	Prospective cohort	1st trimester	Oxazepam, diazepam and lorazepam
LEPPEE2010	893	Croatia	NR	NR	Prospective cohort	Any trimester	Diazepam
OBERLANDER2008	108,288	Canada	30	NR	Prospective cohort	1st trimester	Any benzodiazepine
ORNOY1998	1,989	Israel	30	NR	Prospective cohort	1st trimester	Any benzodiazepine
PASTUSZAK1996	274	Canada	NR	41.6% anxiety disorders; 0.73% benzodiazepine abuse; 8.03% depression; 0.73% drug rehabilitation therapy; 16.06% insomnia; 0.73% OCD; 0.73%	Prospective cohort	1st trimester	Any benzodiazepine

				psychosis; 1.46% seizure			
WIKNER2007	873,879	Sweden	NR	NR	Prospective cohort	Any trimester	Any benzodiazepine
Note. Abbreviations: NR=Not reported							

Table 319. Study information for studies included in the meta-analyses for stimulants

Study ID	Total no. of studies (1); participants	Country	Mean age of participants (years)	Diagnosis of participants	Study design	Timing of exposure	Drugs examined
POTTEGARD2014	2,442	Denmark	NR	NR	Retrospective cohort	2nd trimester	Methylphenidate
Note. Abbreviations: NR=Not reported							

8.4.5 Clinical evidence for adverse events associated with antidepressants (by outcome)

Summary of findings can be found in the tables presented in this section. The associated forest plots can be found in Appendix 19. Data were analysed using meta-analysis. However, outcomes are only presented for analyses with more than one study. In the absence of adequate data, the available evidence was synthesised using narrative methods. Separate analyses were conducted for studies which used a case-control design. It was not possible to conduct sub-group analyses by disordered comparison group as the review was based on existing systematic reviews which did not make this distinction.

Teratogenic harms

The results of the meta-analyses for antidepressants split by individual drug are summarised for congenital malformations (Table 320), major congenital malformation (Table 321), cardiac malformations (Table 322) and septal heart defects (Table 323).

There was some evidence for a statistically significant association between all SSRIs and congenital malformations (p=0.04) with an absolute risk difference of 9 more per 1000. The association between major congenital malformations and all SSRIs was not statistically significant, however the absolute risk difference was 12 more per 1000. Paroxetine was statistically associated with congenital (p=0.05), major congenital (p=0.04) and cardiac (p=0.006) malformations, and fluoxetine with major congenital (p=0.008) and cardiac (p=0.02) malformations with absolute risk differences ranging from 3 to 8 more per 1000. There was some evidence for a statistically significant association between citalopram and escitalopram and ventral septal defects with absolute risk difference of 4 and 9 more per 1000, respectively. It is noteworthy that the association between congenital malformations and TCAs was in favour of the exposed group (absolute risk difference, 20 fewer per 1000), however the baseline rate in the unexposed group was unexpectedly high (137 per 1000).

Course of pregnancy, obstetric and neonatal complications

The results of the meta-analysis for antidepressants split by individual drug are summarised in Table 324. There was some evidence for a statistically significant association between SSRIs in late pregnancy and persistent pulmonary hypertension (p=0.00001), but the actual risk difference was low with only 2 more per 1000 in the SSRI exposed group. However, larger effect sizes were found for an association between any antidepressant and poor neonatal adaptation syndrome, respiratory distress and tremor with absolute risk differences ranging from 34 more to 333 more per 1000. There was also some evidence for greater risk of preterm delivery (17 more per 1000) and miscarriage (12 more per 1000) associated with the SSRI group.

Neurodevelopmental outcomes

There was limited evidence for long-term neurodevelopmental outcomes associated with antidepressants. Risk of autism was not considered in the existing systematic reviews, and additionally searched for. There is evidence that parental mental health problems are themselves associated with autism spectrum disorders in the offspring (Daniels et al., 2008), therefore only studies which used a disorder specific comparison group were included. One study (ELMARROUN2014) found children prenatally exposed to SSRIs had more autistic traits (B= 0.15 [0.08, 0.22]) and were at a higher risk for developing pervasive developmental problems, OR=1.91 (1.31, 3.47) but not affective problems compared with children who were only exposed to depressive symptoms in pregnancy.

Table 320: Summary of findings table for effects of exposure to antidepressants in pregnancy compared with no exposure to antidepressants on congenital malformations

Drug	No of studies (K), Participants (N)	Effect size (OR)	Absolute risk (unexposed)	Absolute risk (exposed)	Absolute risk difference
SSRIs	K=16 N= 2,548,463	1.16 (1.00, 1.35)	43 per 1000	52 per 1000	9 more per 1000
	K[1]=1 N=13,615	1.14 (0.89, 1.47)	N/A	N/A	N/A
TCAs	K=2 N=50,257	0.82 (0.57, 1.18)	137 per 1000	117 per 1000	20 fewer per 1000
Paroxetine	K=8 N=2,372,763	1.20 (1.00, 1.43)	44 per 1000	48 per 1000	4 more per 1000
Citalopram	K=7 N=2,324,723	1.11 (0.91, 1.37)	42 per 1000	59 per 1000	17 more per 1000
Fluoxetine	K=8 N=2,323,821	1.15 (0.96- 1.39)	42 per 1000	42 per 1000	No difference
Sertraline	K=6 N=2,321,611	1.06 (0.80, 1.40)	42 per 1000	39 per 1000	3 fewer per 1000
Fluvoxamine	K=4 N=1,611,180	0.84 (0.48, 1.47)	42 per 1000	29 per 1000	13 fewer per 1000
Escitalopram	K=3 N=1,716,796	1.43 (0.72, 2.87)	41 per 1000	47 per 1000	6 more per 1000

Drug	No of studies (K), Participants (N)	Effect size (OR)	Absolute risk (unexposed)	Absolute risk (exposed)	Absolute risk difference
Venlafaxine	K=2 N=108,652	0.64 (0.32, 1.30)	31 per 1000	20 per 1000	11 fewer per 1000
Note. Abbreviations: OR= Odds ratios; N/A=Not applicable [1] Case-control study design					

Table 321: Summary of findings table for effects of exposure to antidepressants in pregnancy compared with no exposure to antidepressants on major congenital malformations

Drug	No of studies (K), Participants (N)	Effect size (OR)	Absolute risk (unexposed)	Absolute risk (exposed)	Absolute risk difference
Any antidepressant	K[1]=1 N=13,615	1.14 (0.85, 1.53)	N/A	N/A	N/A
All SSRIs	K=11 N=1,250,471	1.15 (0.98, 1.35)	34 per 1000	46 per 1000	12 more per 1000
Paroxetine	K=5 N=1,234,083	1.34 (1.01, 1.78)	34 per 1000	41 per 1000	7 more per 1000
Citalopram	K=5 N=1,233,776	1.11 (0.89, 1.40)	34 per 1000	69 per 1000	35 more per 1000
Fluoxetine	K=6 N=1,234,835	1.27 (1.06, 1.51)	34 per 1000	41 per 1000	7 more per 1000

Setraline	K=4 N=1,231,765	1.15 (0.91, 1.47)	34 per 1000	38 per 1000	4 more per 1000
Fluvoxamine	K=3 N=737,266	0.80 (0.44, 1.46)	35 per 1000	27 per 1000	8 fewer per 1000
Escitalopram	K=2 N=629,048	1.09 (0.67, 1.77)	35 per 1000	39 per 1000	4 more per 1000
Venlafaxine	K=2 N=108,652	0.64 (0.32, 1.30)	31 per 1000	20 per 1000	11 fewer per 1000
Note. Abbreviations: OR= Odds ratios; N/A=Not applicable [1] Case-control study design					

Table 322: Summary of findings table for effects of exposure to antidepressants in pregnancy compared with no exposure to antidepressants on cardiac malformations

Drug	No of studies (K), Participants (N)	Effect size (OR)	Absolute risk (unexposed)	Absolute risk (exposed)	Absolute risk difference
SSRIs	K=10 N=261,216	1.32 (1.01, 1.73)	11 per 1000	13 per 1000	2 more per 1000
TCAs	K=2 N=50,257	0.50 (0.15, 1.66)	24 per 1000	8 per 1000	16 fewer per 1000
Paroxetine	K=7 N=2,371,687	1.46 (1.12, 1.90)	11 per 1000	14 per 1000	3 more per 1000
	K[1]=1 N=1,282	1.53 (0.55, 4.22)	N/A	N/A	N/A

Citalopram	K=5 N=2,323,347	1.41 (0.86, 2.29)	11 per 1000	13 per 1000	2 more per 1000
Fluoxetine	K=6 N=2,322,442	1.58 (1.08, 2.32)	11 per 1000	15 per 1000	4 more per 1000
Setraline	K=5 N=2,320,622	1.29 (0.67, 2.49)	11 per 1000	10 per 1000	1 fewer per 1000
Fluvoxamine	K=2 N=628,847	0.64 (0.16, 2.58)	13 per 1000	8 per 1000	5 fewer per 1000
Escitalopram	K=2 N=842,848	2.54 (0.67, 9.59)	11 per 1000	21 per 1000	10 more per 1000
Venlafaxine	K=1 N=107,570	0.84 (0.12, 5.98)	5 per 1000	4 per 1000	1 fewer per 1000
Note. Abbreviations: OR= Odds ratios; N/A=Not applicable [1] Case-control design					

Table 323: Summary of findings table for effects of exposure to antidepressants compared with no exposure to antidepressants on septal defects (including both atrial septal defects and/or ventral septal defects)

Drug	No of studies (K), Participants (N)	Effect size (OR)	Absolute risk (unexposed)	Absolute risk (exposed)	Absolute risk difference
Both atrial septal defects and/or ventral septal defects					
SSRIs	K=3 N=2,010,497	1.29 (0.97, 1.73)	8 per 1000	11 per 1000	3 more per 1000

Drug	No of studies (K), Participants (N)	Effect size (OR)	Absolute risk (unexposed)	Absolute risk (exposed)	Absolute risk difference
Paroxetine	K=3 N=1,997,822	1.41 (1.01, 1.73)	8 per 1000	12 per 1000	4 more per 1000
Citalopram	K=3 N=2,001,556	1.29 (0.81, 2.04)	8 per 1000	11 per 1000	3 more per 1000
Fluoxetine	K=3 N=1,998,688	1.32 (0.79, 2.23)	8 per 1000	13 per 1000	5 more per 1000
Sertraline	K=3 N=1,998,630	1.23 (0.58, 2.60)	8 per 1000	9 per 1000	1 more per 1000
Fluvoxamine	K=1 N=628,847	0.39 (0.05, 2.75)	11 per 1000	4 per 1000	7 fewer per 1000
Escitalopram	K=1 N=629,048	1.70 (0.85, 3.43)	11 per 1000	18 per 1000	7 more per 1000
Atrial septal defect					
SSRIs	K=2 N=745,528	1.91 (0.85, 3.43)	2 per 1000	2 per 1000	No difference
Paroxetine	K=1 N=629,575	1.52 (0.49, 4.74)	2 per 1000	3 per 1000	1 more per 1000
Citalopram	K=1 N=631,406	1.05 (0.47, 2.35)	2 per 1000	2 per 1000	No difference
Fluoxetine	K=1 N=630,425	1.90 (0.90, 3.99)	2 per 1000	4 per 1000	2 more per 1000

Drug	No of studies (K), Participants (N)	Effect size (OR)	Absolute risk (unexposed)	Absolute risk (exposed)	Absolute risk difference
Setraline	K=1 N=629,476	1.13 (0.28, 4.54)	2 per 1000	2 per 1000	No difference
Ventral septal defects					
SSRIs	K=4 N=745,648	1.39 (0.85, 3.43)	8 per 1000	10 per 1000	2 more per 1000
Paroxetine	K=1 N=629,575	1.19 (0.64, 2.22)	9 per 1000	10 per 1000	1 more per 1000
Citalopram	K=1 N=631,406	1.49 (1.07, 2.07)	9 per 1000	13 per 1000	4 more per 1000
Fluoxetine	K=1 N=630,425	1.65 (1.12, 2.44)	9 per 1000	14 per 1000	5 more per 1000
Setraline	K=1 N=629,471	0.66 (0.27, 1.59)	9 per 1000	6 per 1000	3 fewer per 1000
Fluvoxamine	K=1 N=628,847	0.48 (0.07, 3.40)	9 per 1000	4 per 1000	5 fewer per 1000
Escitalopram	K=1 N=629,048	2.11 (1.05, 4.24)	9 per 1000	18 per 1000	9 more per 1000

Note. Abbreviations: OR=Odds ratios

Table 324: Summary of findings table for effects of exposure to antidepressants compared with no exposure to antidepressants on obstetric and neonatal complications

Harm	Drug	No of studies (K), Participants (N)	Effect size (OR)	Absolute risk (unexposed)	Absolute risk (exposed)	Absolute risk difference
Miscarriage	SSRIs	K=9 N=5,688	1.60 (1.01, 2.53)	12 per 1000	40 per 1000	28 more per 1000
Pre term delivery	SSRIs	K=9 N=225,371	1.38 (0.99, 1.92)	49 per 1000	100 per 1000	51 more per 1000
	TCAs	K=1 N=418	2.01 (0.94, 4.28)	53 per 1000	100 per 1000	47 more per 1000
Poor neonatal adaptation syndrome	Any antidepres	K=6 N=1,954	4.13 (2.14, 7.98)	86 per 1000	366 per 1000	280 more per 1000
	Paroxetine	K=1 N=82	2.23 (0.57, 8.70)	111 per 1000	218 per 1000	107 more per 1000
Persistent pulmonary hypertension	SSRIs	K=1 N=1,599,154	2.51 (1.78, 3.54)	1 per 1000	3 per 1000	2 more per 1000
Respiratory distress	Any antidepressants	K=8 N=754,011	2.07 (1.79, 2.39)	38 per 1000	128 per 1000	90 more per 1000
Tremors	Any antidepressants	K=4 N=482	8.14 (4.23, 15.65)	92 per 1000	444 per 1000	352 more per 1000
Note. Abbreviations: OR=Odds ratios						

8.4.6 Clinical evidence for adverse events associated with antipsychotics (by outcomes)

Summary of findings can be found in the tables presented in this section. The associated forest plots can be found in Appendix 19. Data were analysed using meta-analysis.Due to the limited evidence for antipsychotics, single study outcomes were included in the analyses. In the absence of adequate data, the available evidence was synthesised using narrative methods. Separate analyses were conducted for studies which used a case-control design. Where possible, subgroup analyses were also conducted for studies which used a disorder specific comparison group.

Teratogenic harms

A summary of the meta-analyses for major congenital malformations and congenital malformations is found in Table 325. There was some evidence for a statistically significant association between antipsychotics and congenital and major congenital malformations, with absolute risk differences of 36 more and 13 more per 1000, respectively. When restricting the analysis to one study where the comparison group had a disorder specific comparison group (bipolar disorder), the effect size remained similar, although was no longer statistically significant.

Table 325: Summary of findings table for effects of exposure to antipsychotics compared with no exposure to antipsychotics on congenital and major congenital malformations

Harm	No of studies (K), Participants (N)	Effect size (OR)	Absolute risk (unexposed)	Absolute risk (exposed)	Absolute risk difference
Congenital malformation	K=5 N=1,308,333	1.55 (1.23, 1.95)	38 per 1000	74 per 1000	36 more per 1000
	K [1]=1 N=667	1.81 (0.57, 5.79)	20 per 1000	35 per 1000	15 more per 1000
Major congenital malformation	K=4 N=977,062	1.62 (1.18, 2.22)	31 per 1000	44 per 1000	13 more per 1000

Note. Abbreviations: OR= Odds ratios
[1]Control group consisted of people with bipolar disorder who were not exposed to an antipsychotic

Course of pregnancy, obstetric and neonatal complications

The results of the meta-analysis for course of pregnancy, obstetric and neonatal complications are summarised in Table 326. There was some evidence for a statistically significant association between antipsychotics and gestational diabetes with an absolute risk difference of 19 more per 1000. However the association was

no longer statistically significant and the risk difference reduced to only 1 more per 1000 with a disorder specific comparison, although the sample size was substantially smaller. There was evidence for a significant association between antipsychotics and small for gestational age and low birthweight babies, with large absolute risk differences. However when the control group had a psychiatric diagnosis, the association for small for gestational age was no longer statistically significant and the risk difference reduced. There was evidence for a statistically significant association with preterm delivery and caesarean section with large absolute risk differences of 57 and 103 more per 1000, respectively.

Table 326: Summary of findings table for effects of exposure to antipsychotics compared with no exposure to antipsychotics on course of pregnancy, obstetric and neonatal complications

Harm	No of studies (K), Participants (N)	Effect size	Absolute risk (unexposed)	Absolute risk (exposed)	Absolute risk difference
Gestational diabetes	K=3 N=1,318,376	OR=2.32 (1.53, 3.52)	11 per 1000	30 per 1000	19 more per 1000
	K [1]=1 N=874	OR=1.04 (0.37, 2.89)	18 per 1000	19 per 1000	1 more per 1000
Small for gestational age	K=7 N=944,783	OR=2.30 (1.76, 3.01)	22 per 1000	111 per 1000	89 more per 1000
	K [1]= 2 N=1,566	OR=1.15 (0.82, 1.62)	110 per 1000	119 per 1000	9 more per 1000
Large for gestational age	K=6 N=1,001,085	OR=0.82 (0.65, 1.03)	62 per 1000	56 per 1000	6 fewer per 1000
	K [1]=2 N=1,566	OR=0.82 (0.52, 1.28)	62 per 1000	50 per 1000	12 fewer per 1000
Low birth weight (< 2500 g)	K=2 N=943,994	OR=2.15 (1.60, 2.89)	33 per 1000	80 per 1000	47 more per 1000
	K[2]=1 N=152	OR=5.61 (1.19, 26.52)	N/A	N/A	N/A
Birth weight	K=4 N=624	SMD=-0.02 (-0.18, 0.13)	N/A	N/A	N/A
	K[1]=1 N=32	SMD=-0.38 (-1.09, 0.32)	N/A	N/A	N/A

	K[2]=1 N=152	SMD=-0.27 (-0.59, 0.05)	N/A	N/A	N/A
Preterm delivery	K=8 N=951,825	OR=1.81 (1.39, 2.36)	51 per 1000	108 per 1000	57 per 1000
	K [1]=2 N=1,570	OR=1.58 (0.75, 3.33)	78 per 1000	119 per 1000	41 more per 1000
Miscarriage	K=3 N=3,115	OR=1.26 (0.71, 2.24)	82 per 1000	89 per 1000	7 more per 1000
Still birth	K=5 N=1,335,661	OR=1.45 (0.70, 3.01)	4 per 1000	6 per 1000	2 more per 1000
Caesarean delivery	K=4 N=960,951	OR=1.65 (1.40, 1.95)	149 per 1000	252 per 1000	103 more per 1000
	K [1]=1 N=874	OR=1.12 (0.82, 1.55)	235 per 1000	256 per 1000	21 more per 1000
Gestational age at delivery	K=2 N=531	SMD=-0.09 (-0.29, 0.11)	N/A	N/A	N/A

Note. Abbreviations: OR=Odds ratios
[1] Control group consisted of people with a psychiatric diagnosis who were not exposed to an antipsychotic
[2] Case control study design

Neurodevelopmental complications

The results from the meta-analysis for neurodevelopmental outcomes are summarised in Table Table 327. There was no evidence for a statistically significant difference between the exposed and unexposed groups on the mean score or delayed development on the Bayley scales of Infant and Toddler development at 52 weeks from a single study. There was a single study evidence for statistically significant effect in favour of the unexposed group on the Infant Neurological Battery, however the GDG were mindful that the impact of maternal mental health on the long term development of the infant or child was not adjusted for.

Table Table 327: Summary of findings table for effects of exposure to antipsychotics compared with no exposure to antipsychotics on neurodevelopmental outcomes

	No of studies (K), Participants (N)	**Mean score (SMD)**	**Delayed development (OR)**
Bayley Scales of Infant and Toddler development (52 weeks)			
Cognitive functioning	K=1 N=152	-0.26 (-0.58, 0.06)	1.67 (0.52, 5.36)
Language	K=1 N=152	-0.13 (-0.45, 0.19)	1.13 (0.43, 2.95)
Motor function	K=1 N=152	-0.26 (-0.58, 0.06)	1.67 (0.52, 5.36)
Social/emotional functioning	K=1 N=152	-0.21 (-0.53, 0.11)	2.19 (0.78, 6.17)
Adaptive behaviour	K=1 N=152	-0.20 (-0.51, 0.12)	2.15 (0.70, 6.62)
Infant Neurological Battery (6 months)			
Total score	K=1 N= 107	-0.67 (-1.15, -0.19)	N/A
Note. Abbreviations: N/A= Not applicable			

8.4.7 Clinical evidence for adverse events associated with anticonvulsants (carbamazepine, lamotrigine, valproate) (by outcome)

Summary of findings can be found in the tables presented in this section. The associated forest plots can be found in Appendix 19. Data were analysed using meta-

analysis. However, outcomes are only presented for analyses with more than one study. In the absence of adequate data, the available evidence was synthesised using narrative methods. Separate analyses were conducted for studies which used a case-control design. Where possible, subgroup analyses were also conducted for studies which used a disorder specific comparison group, in the majority of cases this was epilepsy.

Teratogenic harms

The results of the meta-analysis for congenital and major congenital malformations are summarised in Table 328 and for specific teratogenic malformations inTable 329. There was some evidence for a statistically significant association between carbamazepine and congenital malformations and major congenital malformations with absolute risk differences of 62 more and 15 more per 1000. This remained significant when performing a sensitivity analysis for studies with a disordered comparison. The results from the meta-analysis suggested an event rate of 3.5% for major malformations, broadly in line with registry data event rates which range from 2.6% to 5.6%. There was no evidence for a statistically significant association between lamotrigine and major congenital malformations. The absolute risk from the meta-analysis suggested an event rate of 2.8% also in line with existing registry data. There was strong evidence for a statistically significant association between valproate and congenital and major congenital malformations, with a risk difference 20 more per thousand (50 more per thousand when using a disordered comparison). The event rate from the meta-analysis suggests a prevalence of 7.7%, broadly in line with registry data which ranges from 6.7% to 9.7%. There was evidence for a statistically significant association with cleft lip and/or palate and valproate (absolute risk difference 11 more per 1000), and for carbamazepine, although the absolute risk difference was low (4 more per 1000). There was also evidence for a statisitacally significant association between neural tube defects and valproare (absolute risk difference 11 more per 1000), which was not found for either carbamazepine or lamotrigine.

Table 328: Summary of findings table for effects of exposure to anticonvulsants compared with no exposure to anticonvulsants on congenital and major congenital malformations

Drug	No of studies (K), Participants (N)	Effect size (OR)	Absolute risk (unexposed)	Absolute risk (exposed)	Absolute risk difference
Major congenital malformations					
CBZ	K=17 N=10,774	1.89 (1.92, 5.46)	20 per 1000	35 per 1000	15 more per 1000
	K[1]=12 N=6,669	1.43 (1.04, 1.96)	24 per 1000	34 per 1000	10 more per 1000
LMG	K=7 N=842,294	1.48 (0.97, 2.27)	24 per 1000	28 per 1000	4 more per 1000
	K[1]=5 N=3,008	1.41 (0.62, 3.21)	23 per 1000	32 per 1000	9 more per 1000
VPA	K=14 N=108,500	3.37 (2.5, 4.53)	55 per 1000	77 per 1000	22 more per 1000
	K[1]=8 N=3,526	2.6 (1.7, 3.97)	23 per 1000	73 per 1000	50 more per 1000
	K[2]=1 N=76,626	1.51 (1.38, 1.65)	N/A	N/A	N/A
Congenital malformations					
CBZ	K=3 N=1,265	2.22 (1, 4.92)	50 per 1000	112 per 1000	62 more per 1000
	K[1]=2 N=699	3.16 (0.72, 13.78)	19 per 1000	100 per 1000	81 more per 1000

VPA	K[1]=3 N=1,857	4.07 (2.41, 6.88)	24 per 1000	109 per 1000	85 more per 1000

Note. Abbreviations: N/A= Not applicable; CMZ=carbamazepine; LMG=Lamotrigine; VPA=Valproate
[1] Control group consisted of people with a disorder (epilepsy) who were not exposed to an anticonvulsant
[2] Case-control study design

Table 329. Summary of findings table for effects of exposure to anticonvulsants compared with no exposure to anticonvulsants on specific teratogenic malformations

Harm	Drug	No of studies (K), Participants (N)	OR	Absolute risk (unexposed)	Absolute risk (exposed)	Absolute risk difference
Neural tube defects	CBZ	K=1 N=207,257	2.42 (0.77, 7.56)	1 per 1000	3 per 1000	2 more per 1000
	LMG	K=1 N=207,786	1.06 (0.26, 4.29)	1 per 1000	1 per 1000	No difference
		K[1]=1	1.20 (0.29, 4.96)	N/A	N/A	N/A
	VPA	K=1 N=206,547	10.41 (3.85, 28.13)	1 per 1000	12 per 1000	11 more per1000
Cleft lip and/or palate	CBZ	K=1 N=207,257	4.41 (1.82, 10.73)	1 per 1000	5 per 1000	4 more per 1000
	LMG	K[2]=2 N=1,046,265	1.99 (0.20, 19.79)	2 per 1000	4 per 1000	2 more per 1000
		K[1]=2 N=93,641	1.55 (0.33, 7.38)	N/A	N/A	N/A

	VPA	K=1 N=206,547	11.38 (4.21, 30.77)	1 per 1000	12 per 1000	11 more per 1000

Note. Abbreviations: N/A= Not applicable; CMZ=carbamazepine; LMG=Lamotrigine; VPA=Valproate
[1]Case-control study design.
[2]For this analysis data for HERNANDEZ-DIAZ2012 and HOLMES2008 have been combined as they used the same comparison group

Course of pregnancy, obstetric and neonatal complications

The results of the meta-analysis for course of pregnancy, obstetric and neonatal complications are summarised in Table 330. There was limited evidence for neonatal and obstetric complications, however the data suggested no statistically or clinically significant evidence for an increased risk of still birth or perinatal death with carbamazepine. There was an increased risk of preterm birth and carbamazepine but this was not statistically significant. There was limited evidence for neonatal and obstetric complications associated with lamotrigine, but available data suggests there does not appear to be any increased risks. There was some evidence for increase in preterm birth for valproate, although not statistically significant.

Table 330: Summary of findings table for effects of exposure to anticonvulsants compared with no exposure to anticonvulsants on course of pregnancy, obstetric and neonatal complications

Harm	Drug	No of studies (K), Participants (N)	Effect size	Absolute risk (unexposed)	Absolute risk (exposed)	Absolute risk difference
Admission to neonatal care	CBZ	K=1 N=274	1.23 (0.95, 1.59)	89 per 1000	107 per 1000	18 more per 1000
	LMG	K=1 N=1,997	2.25 (1.59, 3.17)	89 per 1000	180 per 1000	91 more per 1000
	VPA	K=1 N=2,479	2.41 (1.89, 3.08)	89 per 1000	191 per 1000	102 more per 1000
Still birth/perinatal death	CBZ	K=2 N=3,202	0.79 (0.12, 5.31)	9 per 1000	9 per 1000	No difference
	LMG	K=1 N=1,973	0.49 (0.03, 8.42)	6 per 10000/	0 per 1000	N/A
	VPA	K=2 N=3,975	1.93 (0.79, 4.7)	4 per 1000	9 per 1000	5 more per 1000
Preterm birth	CBZ	K=2 N=3,202	1.65 (0.64-4.22)	45 per 1000	56 per 1000	11 more per 1000
	LMG	K=1 N=1,973	0.98 (0.47, 2.05)	47 per 1000	46 per 1000	1 fewer per 1000
	VPA	K=2 N=3,804	1.31 (0.94, 1.83)	52 per 1000	62 per 1000	10 more per 1000
Birth-weight	CMZ	K=2 N=461	-0.30 (-0.50, -0.11)	N/A	N/A	N/A

	VPA	K=2 N=2,165	-1.57.58 (-220.12- -95.05)	N/A	N/A	N/A
Note. Abbreviations: N/A= Not applicable; CMZ=carbamazepine; LMG=Lamotrigine; VPA=Valproate						

Neurodevelopmental outcomes

The results of the meta-analysis for neurodevelopmental outcomes are summarised in Table 331. The data did not suggest evidence for an increased risk of longer-term neurodevelopmental complications with carbamazepine or lamotrigine. However, there was evidence for a statistically significant association with valproate and low IQ (particularly verbal IQ), and also with autism at 9 year follow-up.

Table 331: Summary of findings table for effects of exposure to anticonvulsants compared with no exposure to anticonvulsants on neurodevelopmental outcomes

Harm	Drug	No of studies (K), Participants (N)	Effect size	Absolute risk (unexposed)	Absolute risk (exposed)	Absolute risk difference
Full scale IQ	CBZ	K[1]=4 N=377	-3.80 (-16.81, -0.80)	N/A	N/A	N/A
	LMG	K=1 N=93	– 3.15 (-7.87, -1.57)	N/A	N/A	N/A
	VPA	K[1]=4 N=286	-5.06 (-8.42, -1.70)	N/A	N/A	N/A
Verbal IQ	CBZ	K[1]=3 N=289	1.47 (-2.42, 5.36)	N/A	N/A	N/A
	LMG	K=1 N=93	-2.49 (-7.88, 2.90)	N/A	N/A	N/A
	VPA	K[1]=4 N=286	-6.83 (-10.51, -2.15)	N/A	N/A	N/A
Performance IQ	CBZ	K=3 N=289	0.07 (-0.20, 0.34)	N/A	N/A	N/A
	LMG	K=1 N=93	-0.33 (-0.74, 0.08)	N/A	N/A	N/A
	VPA	K=4 N=286	-0.25 (-0.67, 0.17)	N/A	N/A	N/A
Motor development	CBZ	K=2 N=221	2.37 (-3.65, 8.38)	N/A	N/A	N/A

	LMG	K=1 N=92	-0.06 (-0.48, 0.35)	N/A	N/A	N/A
	VPA	N=2 K=184	-0.48 (-0.85, -0.10)	N/A	N/A	N/A
Autism						
Autism checklist (78-week follow-up)	CBZ	K=1 N=262	0.79 (0.22, 2.8)	90 per 1000	73 per 1000	17 fewer per 1000
	LMG	K=1 N=286	1.83 (0.81, 4.13)	90 per 1000	154 per 1000	64 more per 1000
	VPA	K=1 N=246	0.87 (0.19, 3.98)	90 per 1000	80 per 1000	10 fewer per 1000
Autism spectrum disorder (ICD-10) 9-year follow-up	CBZ	K=1 N=655,539	1.25 (0.47, 3.35)	8 per 1000	10 per 1000	2 more per 1000
	LMG	K=1 N=655,394	1.5 (0.75, 3.01)	8 per 1000	12 per 1000	4 more per 1000
	VPA	K=1 N=655,495	3.82 (2.15, 6.80)	8 per 1000	31 per 1000	23 more per 1000

Note. Abbreviations: N/A= Not applicable; CMZ=carbamazepine; LMG=Lamotrigine; VPA=Valproate

[1] Control group consisted of people with a disorder (epilepsy) who were not exposed to an anticonvulsant

8.4.8 Clinical evidence for adverse events associated with lithium (by outcome)

Summary of findings can be found in the tables presented in this section. The associated forest plots can be found in Appendix 19. Data were analysed using meta-analysis. However, outcomes are only presented for analyses with more than one study. In the absence of adequate data, the available evidence was synthesised using narrative methods. Separate analyses were conducted for studies which used a case-control design. It was not possible to conduct subgroup analyses for studies which used a disorder specific comparison group.

Teratogenic harms

The results of the meta-analysis for teratogenic harms are summarised in Table 332. There was limited evidence for lithium due to the small number of studies which provided extractable data. There was some evidence for a statistically significant increase for congenital malformations, however the absolute risk reduction was only 7 more per 1000. Rates of Ebstein's anomaly have previously been associated with lithium exposure. Two studies reporting on Ebstein's anomaly met the inclusion criteria for our review; however, estimates were unstable because of the low number of events, [1 in 20,000 in the general population (Cohen et al.,

1994)]. This was similarly found in a recent systematic review of lithium safety which analysed six case-control studies (N=264) and measured the association between Ebstein's anomaly and lithium (McKnight et al., 2012). They found the odds of exposure to lithium did not differ significantly from controls; however, estimates were unstable because of the low number of events.

Table 332: Summary of findings table for effects of exposure to lithium compared with no exposure to lithium on teratogenic harms

Harm	No of studies (K), Participants (N)	OR	Absolute risk (unexposed)	Absolute risk (exposed)	Absolute risk difference
Congenital malformations	K=4 N=974,914	2.10 (1.21, 3.64)	45 per 1000	52 per 1000	7 more per 1000
	K=2[1] N=782	2.12 (0.80, 5.61)	22 per 1000	54 per 1000	32 more per 1000
	K=1[2] N=33,244	2.21 (0.67, 7.25)	N/A	N/A	N/A
Heart defects	K=2 N=973,967	1.43 (0.59-3.46)	45 per 1000	58 per 1000	13 more per 1000
Ebstein's Anomaly	K=2 N=3,912	Estimates unstable because of low number of events	N/A	N/A	N/A

Note. Abbrieviations; OR=Odds ratios; N/A= Not applicable
[1]Control group consisted of people with a psychiatric diagnosis
[2]Case control study design

Course of pregnancy, obstetric and neonatal complications

There was insufficient evidence for course of pregnancy, neonatal and obstetric complication outcomes.

Neurodevelopmental outcomes

There was insufficient evidence for neurodevelopmental outcomes.

8.4.9 Clinical evidence for adverse events associated with benzodiazepines and related drugs (by outcome)

Summary of findings can be found in the tables presented in this section. The associated forest plots can be found in Appendix 19. Data were analysed using meta-analysis. However, outcomes are only presented for analyses with more than one

study. In the absence of adequate data, the available evidence was synthesised using narrative methods. There was insufficient data to separate out by individual benzodiazepine or related drug, therefore benzodiazepines were considered under one overall class. Separate analyses were conducted for studies which used a case-control design. It was not possible to conduct subgroup analyses for studies which used a disorder specific comparison group.

Teratogenic harms

The results of the meta-analysis for teratogenic harms are summarised in Table 333. The data did not suggest an increased risk of congenital, major congenital or cardiac malformations and benzodiazepines. Data from one cohort study and two case-control studies did not suggest an association with cleft lip or cleft palate.

Table 333: Summary of findings table for effects of exposure to benzodiazepines in pregnancy compared with no exposure to benzodiazepines on teratogenic harms

Harm	No of studies (K), Participants (N)	OR	Absolute risk (unexposed)	Absolute risk (exposed)	Absolute risk difference
Congenital malformation	K=1 N=875,858	1.13 (0.93, 1.38)	47 per 1000	53 per 1000	6 more per 1000
	K[1]=1 N=78	23.20 (4.29, 125.55)	N/A	N/A	N/A
Major congenital malformation	K=5 N=130429	1.01 (0.81-1.25)	31 per 1000	28 per 1000	3 fewer per 1000
	K[1]=1 N=78	19.95 (4.17, 95.45)	N/A	N/A	N/A
Cleft lip with or without a cleft palate	K=2 N=896,995	0.45 (0.23, 1.89)	3 per 1000	1 per 1000	2 fewer per 1000
	K[1]=2 N=4,568	1.52 (0.58, 4.02)	N/A	N/A	N/A
Cardiac abnormalities	K=5 N = 1007764	1.04 (0.56, 1.90)	12 per 1000	8 per 1000	4 fewer per 1000
Septal heart defects	K=1 N=108,288	1.48 (0.21, 10.65)	1 per 1000	1 per 1000	No difference
Atrioventricular defects	K=1 N=108,288	1.52 (0.49, 4.76)	2 per 1000	3 per 1000	1 more per 1000

*Note.*Abbreviations: OR= Odds ratios; N/A= Not applicacble.
[1]Case control study design

Course of pregnancy, obstetric and neonatal complications

The results of the meta-analysis for course of pregnancy, obstetric and neonatal complications are summarised in Table 334. There was some evidence for an increased risk of caesarean delivery and miscarriage and some evidence of an increased risk of respiratory disorder.

Benzodiazepines: neurodevelopmental outcomes

There was insufficient evidence for neurodevelopmental outcomes.

Table 334: Summary of findings table for effects of exposure to benzodiazepines compared with no exposure to benzodiazepines on course of pregnancy, obstetric and neonatal complications

Harm	No of studies (K), Participants (N)	Effect size	Absolute risk (unexposed)	Absolute risk (exposed)	Absolute risk difference
Gestational age at delivery	K=3 N=1,037	SMD=0.02 (-0.13, 0.16)	N/A	N/A	N/A
Birth weight (g)	K=3 N=1,037	SMD=0.02 (-0.17, 0.21)	N/A	N/A	N/A
Caesarean delivery	K=2 N=876,920	OR=1.52 (1.27, 1.81)	49 per 1000	82 per 1000	33 more per 1000
Miscarriage	K=3 N=1,204	OR=1.83 (1.19, 2.82)	59 per 1000	101 per 1000	42 more per 1000
Instrumental delivery	K=2 N=154	OR=1.14 (0.12, 10.69)	354 per 1000	292 per 1000	62 fewer per 1000
Respiratory disorder	K=2 N=875,904	OR=1.26 (1.04, 1.52)	44 per 1000	55 per 1000	11 more per 1000

Note. Abbreviations: OR= Odds ratios; N/A= Not applicable

8.4.10 Clinical evidence for adverse events associated with stimulants (methylphenidate) (by outcome)

Teratogenic harms

The results of the meta-analysis for teratogenic harms are summarised in Table 335. There was no statistically or clinically meaningful association betweenmethylphenidate and congenital and major congenital malformations.

Course of pregnancy, obstetric and neonatal complications

There was insufficient evidence for course of pregnancy, obstetric and neonatal complication outcomes.

Neurodevelopment outcomes

There was insufficient evidence for neurodevelopmental outcomes.

Table 335: Summary of findings table for effects of exposure to stimulants compared with no exposure to stimulants on course of pregnancy, obstetric and neonatal complications

Harm	No of studies (K), Participants (N)	Effect size (OR)	Absolute risk (unexposed)	Absolute risk (exposed)	Absolute risk difference
Major congenital malformations	K=1 N=1,471	1.02 (0.4-2.59)	39 per 1000	40 per 1000	1 more per 1000
Cardiac malformations	K=1 N=1,471	1.92 (0.56-6.65)	13 per 1000	24 per 1000	11 more per 1000
Note. Abbreviations: OR=Odds ratios					

8.5 PHYSICAL INTERVENTIONS FOR THE PREVENTION OF MENTAL HEALTH PROBLEMS IN PREGNANCY AND THE POSTNATAL PERIOD

8.5.1 Clinical review protocol (prevention)

The review protocol summary, including the review question(s), information about the databases searched, and the eligibility criteria used for this section of the guideline, can be found in Table 336. A complete list of review questions can be found in Appendix 8; further information about the search strategy can be found in Appendix 10; the full review protocols can be found in Appendix 9.

The review strategy was to evaluate the clinical effectiveness of the physical interventions for the prevention of mental health problems in pregnancy and the postnatal period using meta-analysis. However in the absence of adequate data, the available evidence was synthesised using narrative methods. An analysis of all interventions was conducted and graded.

Table 336: Clinical review protocol summary for the review of physical interventions for the prevention of mental heal problems

Component	Description
Review question(s)	RQ 2.1 What is the effectiveness of selective preventative interventions (for women with no risk factors) in reducing the likelihood of developing mental health problems in pregnancy or the postnatal period? RQ 2.2 What is the effectiveness of indicated preventative interventions (for women with identified risk factors present) in reducing the likelihood of developing mental health problems in pregnancy or the postnatal period? RQ 2.3 What strategies should be adopted to minimise potential harm to the women or the fetus/infant of these interventions?
Population	Included Review question 2.1 Women who are pregnant or postnatal (from delivery to the end of the first year). Inclusion is not based on any other baseline risk factors. Review question 2.2 Women who are pregnant or postnatal (from delivery to the end of the first year) who are considered to be 'at risk' of developing mental health problems. Include women: with a history of a mental health problem but who do not meet diagnostic criteria for mental health problems at the current time experiencing major life events with a family history of mental health problems with psychosocial risk factors (for example SES) who have infants with regulatory problems who experienced an operative delivery or traumatic birth who experienced a pre-term delivery (<37 weeks gestation) and/or whose infant had a low birth weight who experienced a miscarriage who are adolescents experiencing intimate partner violence Exclude women: who are currently receiving treatment (psychosocial or pharmacological) for an existing mental health problem (see review of interventions for the treatment of a mental health problem) who are not pregnant or postnatal (up to 1 year postnatal)
Intervention(s)	Included interventions Physical interventions for women with no pre-specified baseline risk factors (other than being pregnant or in the postnatal period) (RQ 2.1) or for women with at least one identified baseline risk factor (RQ 2.2), including:

	Physical activity Massage Acupuncture Excluded Interventions Universal prevention programmes (that is, targeted to the general public or to a whole population group that has not been identified on the basis of increased risk)
Comparison	Review question 2.1 and 2.2 Treatment as usual, enhanced treatment as usual, no treatment, waitlist control Another active prevention intervention
Critical outcomes	Maternal Outcomes Symptom-based Diagnosis of mental health problems Symptomatology (clinician- and self-report) Relapse Service utilisation Hospitalisation for mental health problems Retention in services (assessed through dropout rates as a proxy measure) Experience of care Satisfaction Acceptability of treatment (including dropout as a proxy measure) Quality of life Quality of life measures Functional disability Social functioning Social support Perceived parenting stress Harm Side effects (including dropout because of side effects) Quality of mother-infant interaction and infant care Quality of mother-infant interaction measures (including maternal sensitivity and child responsivity) Establishing or continuing breastfeeding Fetal/Infant outcomes Fetal and infant physical development (including congenital malformations) Side effects Cognitive development of the infant Physical development of the infant Emotional development of the infant Optimal care of infant (for example vaccinations, well-baby check-ups) Prevention of neglect or abuse of the infant Service use Planned (health visitor, vaccinations, well-baby check-ups) Unplanned (emergency department visits, inpatient, urgent or acute care) Social service involvement
Study design	Review question 2.1 and 2.2 Systematic reviews of RCTs Primary RCTs

	Review question 2.3 N/A; GDG consensus-based
Note. Abbreviations: N/A= Not applicable	

8.5.2 Studies considered (prevention: no identified risk factors)[18]

Three RCTs met the eligibility criteria for this review: NORMAN2010 (Norman et al., 2010), ROBLEDO-COLONIA2012 (Robledo-Colonia et al., 2012), SONGOYGARD2012 (Songoygard et al., 2012). All studies were published in peer reviewed journals. In addition seven studies were excluded from the review. Further information about the included and excluded studies can be found in Appendix 18.

All studies included sufficient data to be included in the statistical analysis. Of these, two studies (N=811) involved a comparison between physical activity and treatment as usual and one study (N=135) compared physical activity with psychoeducation (Table 337).

Table 337: Study information for trials included in the meta-analyses of physical interventions for the prevention of mental health problems

	Physical activity versus Treatment as usucal	**Physical activity versus psychoeducation**
Total no. of trials (k); participants (N)	2 (935)	1 (161)
Study ID	(1) ROBLEDO-COLONIA2012 (2) SONGOYGARD2012	NORMAN2010
Country	(1) Columbia (2) Norway	Australia
Mean age of participants (years)	(1) 21 (2) 31	30
Timing of intervention	(1-2) Antenatal	Postnatal
Mode of delivery	(1-2) Physiotherapist	Physical therapist
Format	(1-2) Group	Group
Intensity (number of sessions)	(1) Moderate (3 hourly group session per week) (2) Moderate (1 hourly group session per week and 45 minutes twice a week at home)	Low (1 hour group session per week)
Length of intervention (weeks)	(1) 13 (2) 12	8
Setting	(1) Performed in a spacious, air-conditioned room. (2) Not reported	Hospital
Intervention	(1-2) Exercise classes	Group exercise with their babies

[18] Here and elsewhere in the guideline, each study considered for review is referred to by a study ID in capital letters (primary author and date of study publication, except where a study is in press or only submitted for publication, then a date is not used).

Follow-up1	(1) Post-treatment or first measurement (2) Short-term follow-up	Short-term follow-up
Note. 1Time points: Post-treatment or first measurement; Short-term follow-up (9-16 weeks post-intervention); Intermediate follow-up (17-24 weeks post-intervention); Long-term follow-up (25-103 weeks post-intervention follow-up); Very long-term follow-up (=>104 weeks).		

8.5.3 Clinical evidence for physical interventions (prevention no identified risk factors)

Summary of findings can be found in the tables presented in this section. The full GRADE evidence profiles and associated forest plots can be found in Appendix 22 and Appendix 19, respectively.

Physical activity versus treatment as usual There was low quality, single study (N=74) evidence for a large beneficial preventative effect of physical activity on mean depression scores at the end of the intervention (p=0.0006, Table 338). In addition, there was low quality, single study (N=737) evidence for a large preventative effect of physical activity on depression symptomology (above threshold), p=0.16. However there was very serious imprecision due to the small number of events and the 95% CI included both no effect and the measure of appreciable benefit.

Table 338: Summary of findings tables for the preventative effects of physical interventions on depression outcomes

Physical activity compared with control for preventing depression during pregnancy and the postnatal period					
Patient or population: women who are pregnant or postpartum **Settings:** **Intervention: Physical activity** **Comparison: Control group**					
Outcomes	**Illustrative comparative risks* (95% CI)** **Assumed risk** **Control group**	**Corresponding risk** **Physical activity**	**Relative effect (95% CI)**	**No of Participants (studies)**	**Quality of the evidence (GRADE)**	**Comments**
Depression mean scores (post-treatment, 0-8 weeks) -		The mean depression mean scores (post-treatment, 0-8 weeks) - available case analysis in the intervention groups		74 (1 study)	⊕⊕⊖⊖ low1,2	SMD -0.84 (-1.32 to -0.36)

Available case analysis		was 0.84 standard deviations lower (1.32 to 0.36 lower)			
Above depression threshold (short term follow-up, 9-16 weeks) - Available case analysis	Study population		RR 0.43 (0.13 to 1.41)	737 (1 study)	⊕⊕⊖⊖ low1,2
	24 per 1000	10 per 1000 (3 to 33)			
	Moderate				
	24 per 1000	10 per 1000 (3 to 34)			

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Risk of bias in several domains
2 Optimal information size (for dichotomous outcomes, OIS=300 events; for continuous outcomes, OIS=400 participants) not met.

Physical activity combined with psychoeducation versus psychoeducation

There was no statistically or clinically significant effect of physical activity combined with psychoeducation on mean depression scores (p=0.17) at the end of intervention from low quality, single study (N=135) evidence (Table 339). However there was a trend (p=0.06) towards a preventative beneficial effect at short term follow-up using an ITT (LOCF) analysis, however the effect size failed to reach the threshold for a measure of clinically appreciable benefit.

Table 339: Summary of findings tables for the effects of physical interventions on preventing depression outcomes in women who are pregnant or postpartum

Physical activity and psychoeducation (non-mental health) compared with psychoeducation alone (non-mental health) for preventing depression during pregnancy and the postnatal period

Patient or population: women who are pregnant or postpartum
Settings:
Intervention: Physical activity and psychoeducation (non-mental health)
Comparison: Psychoeducation alone (non-mental health)

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk Psychoeducation alone (non-mental health)	Corresponding risk Physical activity and psychoeducation (non-mental health)				
Depression mean scores- post-treatment (0-8 weeks) - ITT LOCF		The mean depression mean scores- post-treatment (0-8 weeks) – ITT LOCF in the intervention groups was 0.24 standard deviations lower (0.58 lower to 0.1 higher)		135 (1 study)	⊕⊕⊖⊖ low1,2	SMD -0.24 (-0.58 to 0.1)
Depression mean scores- short term follow-up (9-16 weeks) - ITT LOCF		The mean depression mean scores- short term follow-up (9-16 weeks) – ITT LOCF in the intervention groups was 0.33 standard deviations lower (0.67 lower to 0.01 higher)		135 (1 study)	⊕⊕⊖⊖ low1,2	SMD -0.33 (-0.67 to 0.01)

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Optimal information size (for dichotomous outcomes, OIS=300 events; for continuous outcomes, OIS=400 participants) not met.
2 Risk of bias in several domains

8.5.4 Health economics evidence

Systematic literature review

No studies assessing the cost effectiveness of physical interventions for the prevention of mental health problems in pregnancy or the postnatal period were identified by the systematic search of the economic literature undertaken for this guideline. Details on the methods used for the systematic search of the economic literature are described in Chapter 3.

8.5.5 Studies considered: prevention (risk factors identified)

One RCT met the eligibility criteria for this review: HADDAD-RODRIGUES2013 (Haddad-Rodrigues et al., 2013; Table 340). This study compared acupuncture with placebo acupuncture. One study was excluded from the review. Further information about the included and excluded study can be found in Appendix 18.

Table 340: Study information for trials included in the meta-analyses of physical interventions for the prevention of mental health problems

	Acupuncture versus placebo acupuncture
Total no. of trials (k); participants (N)	1 (29)
Study ID	HADDAD-RODRIGUES2013
Country	Brazil
Mean age of participants (years)	27
Timing of intervention	Postnatal
Mode of delivery	Licensed nurse acupuncturist
Format	Individual
Intensity (number of sessions)	Not reported (unclear)
Length of intervention (weeks)	12
Setting	Clinic-primary
Intervention	Acupuncture
Follow-up[1]	Post-treatment
Note. [1] Time points: Post-treatment or first measurement; Short-term follow-up (9-16 weeks post-intervention); Intermediate follow-up (17-24 weeks post-intervention); Long-term follow-up (25-103 weeks post-intervention follow-up); Very long-term follow-up (≥104 weeks).	

8.5.6 Clinical evidence for physical interventions (prevention identified risk factors)

Summary of findings can be found in the tables presented in this section. The full GRADE evidence profiles and associated forest plots can be found in Appendix 22 and Appendix 19, respectively.

Acupuncture versus placebo acupuncture

There was no statistically or clinically significant effect of acupuncture on mean anxiety scores (p=0.14) or cortisol levels (p=1.00) at the end of intervention (Table 341).

Table 341. Summary of findings tables for the effects of acupuncture on preventing anxiety outcomes in women who are pregnant or postpartum

Anxiety: Acupuncture versus control for [health problem]

Patient or population: patients with [health problem]
Settings:
Intervention: Anxiety: Acupuncture versus control

Outcomes	Illustrative comparative risks* (95% CI)		effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk Control	Corresponding risk Anxiety: Acupuncture versus control				
Anxiety mean scores- Post intervention- Available case analysis STAI Follow-up: 12 weeks		The mean anxiety mean scores- post intervention- available case analysis in the intervention groups was **0.56 standard deviations higher** (0.19 lower to 1.3 higher)		29 (1 study)	⊕⊖⊖⊖ **very low**[1,2]	SMD 0.56 (-0.19 to 1.3)
Cortisol mean levels- Post-intervention- Available case analysis Follow-up: 12 weeks		The mean cortisol mean levels- post-intervention- available case analysis in the intervention groups was **0 standard deviations higher** (0.73 lower to 0.73 higher)		29 (1)	very low[1,2]	SMD 0 (-0.73 to 0.73)

*The basis for the **assumed risk** (for example the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% CI) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] High risk of bias in several domains
[2] Total population size is less than 400 (a threshold rule-of-thumb) and 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5 or RR 0.75/1.25)

8.6 PHYSICAL INTERVENTIONS FOR THE TREATMENT OF MENTAL HEALTH PROBLEMS IN PREGNANCY AND THE POSTNATAL PERIOD

8.6.1 Clinical review protocol (treatment)

The review protocol summary, including the review question(s), information about the databases searched, and the eligibility criteria used for this section of the

guideline, can be found in Table 342. A complete list of review questions can be found in Appendix 8; further information about the search strategy can be found in Appendix 10; the full review protocols can be found in Appendix 9.

The review strategy was to evaluate the clinical effectiveness of the physical interventions for the prevention of mental health problems in pregnancy and the postnatal period using meta-analysis. However in the absence of adequate data, the available evidence was synthesised using narrative methods. An analysis of all interventions was conducted and graded.

Table 342: Clinical review protocol summary for the review of physical interventions for the treatment of mental health problems

Component	Description
Review question(s)	RQ 4.2 For women with mental health problems in pregnancy or the postnatal period, what are the benefits and/or potential harms of **physical** interventions to treat mental health problems?
Population	Included Women who have mental health problems in pregnancy and the postnatal period (from delivery to the end of the first year). Include: Women with subthreshold symptoms (but no formal diagnosis of a mental health problem) Women with a formal diagnosis of mild, moderate and severe disorders Exclude: Women who are not pregnant or postnatal (up to 1 year postnatal)
Intervention(s)	Physical interventions, including: Physical activity Massage Acupuncture
Comparison	Treatment as usual, enhanced treatment as usual, no treatment, waitlist control Another active intervention
Critical outcomes	Maternal Outcomes Symptom-based Diagnosis of mental health problem Symptomatology Relapse Use of drugs/alcohol Service utilisation Hospitalisation Retention in services (assessed through dropout rates as a proxy measure) Health service utilisation (for instance, use of psychiatric services) Experience of care Satisfaction (validated measures only, specific items will not be analysed) Acceptability of treatment (assessed through questioning or through including dropout as a proxy measure) Quality of life Quality of life measures

	Functional disability Social functioning Social support Self-esteem Perceived parenting stress Maternal confidence Preservation of rights Harm Side effects (including dropout because of side effects) Maternal mortality and serious morbidity including self-harm and suicide attempts Quality of mother-infant interaction Quality of mother-infant interaction (including maternal sensitivity and child responsivity) Maternal attitude towards motherhood Establishing or continuing breastfeeding Infant outcomes (no restriction on length of follow-up) Fetal and infant physical development (including congenital malformations) Side effects (especially of pharmacological interventions for the fetus and for the infant if breastfeeding) Apgar score Birth weight Admission to neonatal intensive care unit Cognitive development of the infant Emotional development of the infant Physical development of the infant Prevention of neglect or abuse of the infant Optimal care of infant (for example vaccinations, well-baby check-ups) Foetal/infant mortality Foetal/infant morbidity Service use Planned (health visitor, vaccinations, well-baby check-ups) Unplanned (emergency department visits, inpatient, urgent or acute care) Social service involvement
Study design	Systematic reviews of RCTs Primary RCTs For protocols for women following stillbirth, cohort studies were included
Note.	

8.6.2 Studies considered[19] (treatment)

In total, ten RCTs met the eligibility criteria for this review: ARMSTRONG2004 (Armstrong et al., 2004), CHUNG2012 (Chung et al. 2012), DALEY2008 (Daley et al., 2008), DALEY2014 (Daley et al., 2014), FIELD2013B (Field et al., 2013b,

[19] Here and elsewhere in the guideline, each study considered for review is referred to by a study ID in capital letters (primary author and date of study publication, except where a study is in press or only submitted for publication, then a date is not used).

MANBER2004 (Manber et al., 2004), MANBER2010 (Manber et al., 2010), O'HIGGINS2008 (O'Higgins et al., 2008), ONOZAWA2001 (Onozawa et al., 2001), WIRZ-JUSTICE2011 (Wirz-Justice et al., 2011). All were published in peer-reviewed journals between 2001 and 2012. One study (DALEY2014) was in press at the time of the review. In addition, ten studies were excluded from the review. Further information about the included and excluded studies can be found in Appendix 18.

There were two studies which compared physical activity and treatment as usual, and one study that compared physical activity with mutual support (Table 343).

There was one study involved a comparison between acupuncture and massage and one study between depression-specific acupuncture compared with non-depression specific acupuncture, one study which involved a comparison between electro-acupuncture and non-invasive sham acupuncture, one study that compared massage with support and one study compared massage combined with support compared with support alone (Table 344).

Finally, one study compared bright light therapy with placebo (Table 345).

Table 343: Study information table for trials included in the treatment meta-analysis of physical activity interventions versus any alternative treatment intervention

	Physical activity compared with treatment as usual	**Physical activity compared with mutual support**
Total no. of trials (k); participants (N)	3 (191)	1 (24)
Study ID	(1) DALEY2008 (2) DALEY2014 (3) FIELD2013B	ARMSTRONG2004
Country	(1-2) UK (3)US	Australia
Mean age of participants (years)	(1) NR (2) 30 (3) 27	NR
Baseline diagnostic status	(1) Symptoms of depression (EPDS score >12) (2) Diagnosis of major depressive disorder (ICD-10) (3) SCID for DSM-IV	Symptoms of depression (EPDS score ≥12)
Timing of intervention	(1-2) Postnatal (3)Pregnancy	Postnatal
Mode of delivery	(1) Trained researcher (2) Physical activity facilitator (3) Trained yoga instructor	Facilitators (nurse/social worker)
Format	(1) Individual (2) Group and individual (3) Group	Group

Intensity (number of sessions)	(1) Low (two hourly sessions and follow-up support calls for 7 weeks). (2) Low (two hourly sessions and two telephone support calls) (3) Low (20 min sessions for 12 weeks)	Moderate (two 40 min sessions per week [plus one solo session])
Length of intervention (weeks)	(1) 12 (2) 26 (3) 12	12
Time points[1]	(1) Post-treatment (2) Post-treatment and long-term follow-up (3) Post-treatment	Post-treatment
Setting	(1) Home (2) Not reported (3) Not reported	Not reported
Intervention	(1-2) Exercise consultations Tai-chi/yoga	Pram walking exercise programme
Follow-up[1]	(1) Post-treatment (2) Post-treatment; Long-term follow-up (3) Post-treatment	Post-treatment
Note. Abbreviations: NR= Not reported [1]Time points: Post-treatment or first measurement; Short-term follow-up (9-16 weeks post-intervention); Intermediate follow-up (17-24 weeks post-intervention); Long-term follow-up (25-103 weeks post-intervention follow-up); Very long-term follow-up (=>104 weeks).		

Table 344: Study information table for trials included in treatment meta-analysis of any acupuncture or massage interventions versus any alternative treatment intervention

	Acupuncture[2] compared with massage	Depression specific acupuncture compared with non-depression specific acupuncture	Electro-acupuncture compared with non-invasive sham acupuncture	Massage compared with support	Massage combined with support compared with support alone
Total no. of trials (k); participants (N)	2 (210)	2 (210)	1 (20)	1 (62)	1 (34)
Study ID	(1) MANBER2004 (2) MANBER2010	(1) MANBER2004 (2) MANBER2010	CHUNG2012	O'HIGGINS2008	ONOZAWA2001
Country	(1-2) USA	(1-2) USA	China	UK	UK
Mean age of participants (years)	(1-2) 33	(1-2) 33	35	NR	34
Baseline diagnostic status	(1-2) Diagnosis of major depressive disorder (DSM-IV)	(1-2) Diagnosis of major depressive disorder (DSM-IV)	Major depressive episode (measure not reported)	Symptoms of depression (EPDS score >12)	Symptoms of depression (EPDS score >12)
Timing of intervention	(1-2) Antenatal	(1-2) Antenatal	Postnatal	Postnatal	Postnatal
Mode of delivery	(1-2) Masseur	(1-2) Acupuncturist	Acupuncturist	Trained infant masseurs	Trained instructor
Format	(1-2) Individual	(1-2) Individual	Individual	Group	Group
Intensity (number of sessions)	(1-2) Moderate (12 half an hour sessions)	(1-2) Moderate (12 half an hour sessions)	Moderate (2 sessions a week)	Moderate (6 sessions overall)	Low (1 hour session a week)
Length of intervention (weeks)	(1-2) 8	(1-2) 8	4	No defined start or end session	5
Setting	(1-2) Not reported	(1-2) Not reported	Clinic	Not reported	Hospital clinic
Intervention	(1-2) Massage	(1-2) Depression specific acupuncture	Electroacupuncture	Infant massage classes	Infant massage group and a social support group
Time points[1]	(1-2) : Post-treatment or first measurement; Short term follow-up	(1-2) Post-treatment or first measurement; Short term follow-up	Post-treatment or first measurement;	Post-treatment or first measurement; Long-term follow-up	Post-treatment or first measurement

Note. Abbreviations: NR= Not reported

[1]Time points; Short-term follow-up (9-16 weeks post-intervention); Intermediate follow-up (17-24 weeks post-intervention); Long-term follow-up (25-103 weeks post-intervention follow-up); Very long-term follow-up (=>104 weeks).

[2]Data from the depression specific and non-depression specific acupuncture have been combined from MANBER2004

Table 345: Study information table for trials included in treatment meta-analysis of bright-light therapy versus placebo

	Bright light compared with Placebo
Total no. of trials (k); participants (N)	1 (46)
Study ID	WIRZ-JUSTICE2011
Country	Switzerland
Mean age of participants (years)	32
Baseline diagnostic status	Diagnosis of major depressive disorder (DSM-IV)
Timing of intervention	Antenatal
Mode of delivery	Light box at home
Format	Individual
Intensity (number of sessions)	High (1 hour a day)
Length of intervention (weeks)	5
Setting	Home
Intervention	Bright light therapy
Time points[1]	Post-treatment or first measurement
[1]Time points: Post-treatment or first measurement; Short-term follow-up (9-16 weeks post-intervention); Intermediate follow-up (17-24 weeks post-intervention); Long-term follow-up (25-103 weeks post-intervention follow-up); Very long-term follow-up (=>104 weeks).	

8.6.3 Clinical evidence for physical interventions (treatment)

Summary of findings can be found in the tables presented in this section. The full GRADE evidence profiles and associated forest plots can be found in Appendix 22 and Appendix 19, respectively.

Response outcomes (by intervention)

Acupuncture versus massage

There was no statistically or clinically significant difference in effect for acupuncture compared with massage on depression outcomes at post-treatment (p=0.27, Table 346).

Table 346: Summary of findings tables for treatment effects of acupuncture versus massage on response outcomes

Acupuncture compared with massage for depression in pregnancy and the postnatal period						
Patient or population: patients with depression in pregnancy and the postnatal period **Settings:** **Intervention: Acupuncture** **Comparison: Massage**						
Outcomes	**Illustrative comparative risks* (95% CI)**		**Relative effect (95% CI)**	**No of Participants (studies)**	**Quality of the evidence (GRADE)**	**Comments**
	Assumed risk **Massage**	**Corresponding risk** **Acupuncture**				
Non-response-post-treatment (0-8 weeks) -	Study population		RR 0.8 (0.54 to 1.19)	188 (2)	⊕⊖⊖⊖ very low1,2	
	442 per 1000 (298 to 657)	355 per 1000 (224 to 562)				
	Moderate					
	466 per 1000 (315 to 694)	379 per 1000 (239 to 600)				

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Risk of bias in several domains
2 Optimal information size (for dichotomous outcomes, OIS=300 events; for continuous outcomes, OIS=400 participants) not met

Depression specific acupuncture versus non-depression specific acupuncture

There was very low quality evidence from two studies (N=121) for a moderate beneficial effect of depression-specific acupuncture post-treatment (p=0.009, Table 347). However, the confidence in this estimate was very low due to serious imprecision (small number of events) and risk of bias in several domains.

Table 347: summary of findings tables for effects of depression-specific acupuncture versus non-depression-specific acupuncture on response outcomes

Depression specific acupuncture compared with non-depression specific acupuncture for depression in pregnancy and the postnatal period

Patient or population: patients with depression in pregnancy and the postnatal period
Settings:
Intervention: Depression specific acupuncture
Comparison: Non-depression specific acupuncture

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Non-depression specific acupuncture	Depression specific acupuncture				
Non-response – [1]HRSD >=14 and >=50% reduction from baseline	Study population		RR 0.59 (0.4 to 0.88)	121 (2 studies)	⊕⊖⊖⊖ very low1,2	
	593 per 1000	350 per 1000 (237 to 522)				
	Moderate					
	576 per 1000	340 per 1000 (230 to 507)				

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Risk of bias in several domains
2 Optimal information size (for dichotomous outcomes, OIS=300 events; for continuous outcomes, OIS=400 participants) not met.

Bright light therapy versus placebo

The results for response to treatment for bright light therapy were inconsistent. There was very low quality, single study (N=27) evidence for a large beneficial effect on response (p=0.06) and remission (p=0.10) to treatment at the end of intervention using an available case-analysis, however the effect was not statistically significant (Table 348). Moreover, the confidence in this estimate was very low due to serious

imprecision (small number of events and the 95% CI included both no effect and measure of appreciable benefit) and risk of bias in several domains.

Table 348: Summary of findings tables for treatment effects of bright light therapy versus placebo on response outcomes

Bright light therapy compared with placebo for depression in pregnancy and the postnatal period

Patient or population: patients with Depression in pregnancy and the postnatal period
Settings:
Intervention: Bright light therapy
Comparison: Placebo

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk Placebo	Corresponding risk Bright light therapy				
Response at post-treatment - Non-response (Structured Interview Guide for the Hamilton Depression Rating Scale--Seasonal Affective Disorder -Atypical depression supplement <50% improvement) - available case analysis	Study population 636 per 1000	248 per 1000 (95 to 655)	RR 0.39 (0.15 to 1.03)	27 (1 study)	⊕⊖⊖⊖ very low1,2	
	Moderate 636 per 1000	248 per 1000 (95 to 655)				
Remission at post-treatment - Non-remission (HDRS <50% improvement to final score >8) - available case analysis	Study population 636 per 1000	312 per 1000 (134 to 732)	RR 0.49 (0.21 to 1.15)	27 (1 study)	⊕⊖⊖⊖ very low1,2	
	Moderate 636 per 1000	312 per 1000 (134 to 731)				

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Risk of bias in several domains
2 Optimal information size (for dichotomous outcomes, OIS=300 events; for continuous outcomes, OIS=400 participants) not met.

Depression outcomes (by intervention)

Physical activity versus treatment as usual

There was no evidence for a statistically or clinically meaningful effect of physical activity on mean depression scores at the end of intervention (p= 0.11), although the effect favoured physical activity compared with control

Table 349: Summary of findings tables for treatment effects of physical interventions versus treatment as usual on depression outcomes

Depression: Physical activity compared to control for depression in pregnancy and the postnatal period

Patient or population: patients with
Settings:
Intervention: Depression: Physical activity
Comparison: control

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Depression: Physical activity				
Depression mean scores- Post intervention, first available endpoint data - available case analysis Follow-up: 12-26 weeks		The mean depression mean scores- post intervention, first available endpoint data - available case analysis in the intervention groups was 0.23 standard deviations lower (0.52 lower to 0.05 higher)		191 (3 studies)	⊕⊕⊖⊖ low1,2	SMD -0.23 (-0.52 to 0.05)

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Unclear risk of bias in several domains
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5) or RR 0.75/1.25 and optimal information size (400 participants) not met

Physical activity versus mutual support

There was very low quality, single study (N=19) evidence for a large beneficial effect of physical activity compared with mutual support on mean depression scores at post-treatment (p=0.04) and at short-term follow-up (p=0.03, Table 350). However, the confidence in this estimate was very low due to serious imprecision (very small population size) and risk of bias in several domains.

Table 350: Summary of findings tables for treatment effects of physical interventions versus mutual support on depression outcomes

Physical activity compared with mutual support for depression in pregnancy and the postnatal period

Patient or population: patients with depression in pregnancy and the postnatal period
Settings:
Intervention: Physical activity
Comparison: Mutual support

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Mutual support	 Corresponding risk Physical activity	Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
Depression mean scores (post-treatment, 0-9 weeks) - available case analysis		The mean depression mean scores (post-treatment, 0-9 weeks) - available case analysis in the intervention groups was 1.05 standard deviations lower (2.02 to 0.07 lower)		19 (1 study)	⊕⊖⊖⊖ very low1,2	SMD -1.05 (-2.02 to -0.07)
Depression mean scores (short term follow-up, 9-16 weeks) - available case analysis		The mean depression mean scores (short term follow-up, 9-16 weeks) - available case analysis in the intervention groups		19 (1 study)	⊕⊖⊖⊖ very low1,2	SMD -1.09 (-2.07 to -0.11)

	was 1.09 standard deviations lower (2.07 to 0.11 lower)

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Risk of bias in several domains
2 Optimal information size (for dichotomous outcomes, OIS=300 events; for continuous outcomes, OIS=400 participants) not met

Acupuncture versus massage

There was no statistically or clinically significant difference in effect for acupuncture (depression and non-depression specific acupuncture combined) compared with massage on mean depression scores at post-treatment or short term follow-up (Table 351). There was very low quality evidence for a moderate beneficial effect of acupuncture compared with massage on depression diagnosis at short term follow-up (p=-0.31), but this was not statistically significant and the confidence in the estimate of the effect is low due to very serious imprecision (low number of events and the 95% CIs included both no effect and a measure of appreciable benefit.

Table 351: Summary of findings tables for treatment effects of acupuncture versus massage on depression outcomes

Acupuncture compared with massage for depression in pregnancy and the postnatal period

Patient or population: patients with depression in pregnancy and the postnatal period
Settings:
Intervention: Acupuncture
Comparison: Massage

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Massage	Corresponding risk Acupuncture	Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
Depression mean scores- post-treatment (0-8 weeks)- available case analysis		The mean depression mean scores- post-treatment (0-8 weeks) in the intervention groups was 0.19 standard deviations higher (0.47 lower to 0.85 higher)		54 (1)	⊕⊖⊖⊖ very low1,2	SMD 0.19 (-0.47 to 0.85)
Depression mean scores- short term follow-up (9-16 weeks) available case analysis		The mean depression mean scores- short term follow-up (9-16 weeks) in the intervention groups was 0.16 standard deviations lower (0.77 lower to 0.45 higher)		49 (1)	⊕⊖⊖⊖ very low1,2	SMD -0.16 (-0.77 to 0.45)
Above depression threshold (DSM-IV)- short term follow-up (9-16 weeks) - available case analysis	Study population 286 per 1000 Moderate 286 per 1000	71 per 1000 (9 to 660) 72 per 1000 (9 to 661)	RR 0.44 (0.09 to 2.13)	46 (1)	⊕⊖⊖⊖ very low1,2	

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of

effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Risk of bias in several domains
3 Optimal information size (for dichotomous outcomes, OIS=300 events; for continuous outcomes, OIS=400 participants) not met

Depression-specific acupuncture versus non-depression specific acupuncture

There was no statistically or clinically significant difference between depression-specific acupuncture and non-depression specific acupuncture on mean depression scores at post-treatment or short term follow-up (Table 352). However there was very low quality, single study (n=35) evidence for a moderate to large effect in the favour of depression-specific acupuncture on depression diagnosis at the end of intervention (p=0.33) and at short term follow-up (p=0.71), however these effects were not statistically significant and confidence in this estimate is very low due to very serious imprecision (very small number of events and the 95% CI crosses both the line of no effect and measure of appreciable benefit or harm).

Table 352: Summary of findings tables for treatment effects of depression-specific acupuncture versus non-depression specific acupuncture on depression outcomes

Depression specific acupuncture compared with non-depression specific acupuncture for depression in pregnancy and the postnatal period

Patient or population: patients with depression in pregnancy and the postnatal period
Settings:
Intervention: Depression specific acupuncture
Comparison: Non-depression specific acupuncture

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Non-depression specific acupuncture	Depression specific acupuncture				
Depression mean scores- post-treatment		The mean depression mean scores- post-treatment (0-8 weeks) in the intervention		35 (1 study)	⊕⊖⊖⊖ very low1,2	SMD -0.38 (-1.06 to 0.29)

(0-8 weeks)- available case		groups was 0.38 standard deviations lower (1.06 lower to 0.29 higher)				
Depression mean scores - short term follow-up (9-16 weeks) available case		The mean depression mean scores - short term follow-up (9-16 weeks) in the intervention groups was 0.12 standard deviations lower (0.82 lower to 0.57 higher)		32 (1 study)	⊕⊖⊖⊖ very low1,2	SMD -0.12 (-0.82 to 0.57)
Above depression threshold (DSM-IV)- post-treatment (0-8 weeks) available case	Study population		RR 0.47 (0.11 to 2.13)	35 (1 study)	⊕⊖⊖⊖ very low1,2	
	263 per 1000	124 per 1000 (29 to 561)				
	Moderate					
	263 per 1000	124 per 1000 (29 to 560)				
Above depression threshold (DSM-IV)- short term follow-up (9-16 weeks) available case	Study population		RR 0.64 (0.06 to 6.39)	32 (1 study)	⊕⊖⊖⊖ very low1,2	
	111 per 1000	71 per 1000 (7 to 710)				
	Moderate					
	111 per 1000	71 per 1000 (7 to 709)				

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Risk of bias in several domains
2 Optimal information size (for dichotomous outcomes, OIS=300 events; for continuous outcomes, OIS=400 participants) not met.

Electroacupuncture versus non-invasive sham acupuncture

There was no statistically or clinically significant effect for electroacupuncture on mean depression scores at post-treatment (p=0.65, Table 353).

Table 353: Summary of findings tables for treatment effects of electroacupuncture versus non-invasive sham acupuncture on depression outcomes

Electroacupuncture compared with non-invasive sham acupuncture for depression in pregnancy and the postnatal period

Patient or population: patients with Depression in pregnancy and the postnatal period
Settings:
Intervention: Electroacupuncture
Comparison: Non-invasive sham acupuncture

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Non-invasive sham acupuncture	 Corresponding risk Electroacupuncture	Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
Depression mean scores- post-treatment (0-8 weeks) - available case analysis		The mean depression mean scores- post-treatment (0-8 weeks) - available case analysis in the intervention groups was 0.21 standard deviations lower (1.09 lower to 0.67 higher)		20 (1 study)	⊕⊖⊖⊖ very low1,2	SMD -0.21 (-1.09 to 0.67)

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Risk of bias in several domains
2 Optimal information size (for dichotomous outcomes, OIS=300 events; for continuous outcomes, OIS=400 participants) not met.

Massage combined with support versus support

There was very low quality, single study (N=25) evidence for a large beneficial effect of massage combined with support compared with support alone on mean

depression scores post-treatment using an available case analysis (p=0.005, Table 354). However the confidence in this estimate was very low due to serious imprecision (very small population size) and there was a risk of bias in several domains.

Table 354: Summary of findings tables for treatment effects of electroacupuncture versus non-invasive sham acupuncture on depression outcomes

Massage and a support group compared with a support group alone for depression in pregnancy and the postnatal period

Patient or population: patients with depression in pregnancy and the postnatal period
Settings:
Intervention: Massage + support group
Comparison: Support group alone

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Support group alone	 Corresponding risk Massage + support group	Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
Depression mean scores- post-treatment (0-8 weeks) - Available case analysis		The mean depression mean scores- post-treatment (0-8 weeks) - available case analysis in the intervention groups was 1.23 standard deviations lower (2.1 to 0.36 lower)		25 (1 study)	⊕⊖⊖⊖ very low1,2	SMD -1.23 (-2.1 to -0.36)

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Risk of bias in several domains
2 Optimal information size (for dichotomous outcomes, OIS=300 events; for continuous outcomes, OIS=400 participants) not met.

Massage versus support

There was no statistically or clinically significant effect of massage compared with support on mean depression scores at post-treatment (p=0.20) or short term follow-up (p=0.70, Table 355).

Table 355: Summary of findings tables for treatment effects of massage versus support on depression outcomes

Massage compared with support for depression in pregnancy and the postnatal period

Patient or population: patients with depression in pregnancy and the postnatal period
Settings:
Intervention: Massage
Comparison: Support group

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Support group	Corresponding risk Massage	Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
Depression mean scores- post-treatment (0-8 weeks) - available case		The mean depression mean scores- post-treatment (0-8 weeks) - available case in the intervention groups was 0.33 standard deviations lower (0.83 lower to 0.18 higher)		61 (1 study)	⊕⊖⊖⊖ very low1,2	SMD -0.33 (-0.83 to 0.18)
Depression mean scores- long term follow-up (>24 weeks) - available case analysis		The mean depression mean scores- long term follow-up (>24 weeks) - available case analysis in the intervention groups was 0.11 standard deviations lower (0.68 lower to 0.46 higher)		48 (1 study)	⊕⊖⊖⊖ very low1,2	SMD -0.11 (-0.68 to 0.46)

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the

estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Risk of bias in several domains
2 Optimal information size (for dichotomous outcomes, OIS=300 events; for continuous outcomes, OIS=400 participants) not met.

Bright light therapy versus placebo

Although there was a trend towards a beneficial effect of bright light therapy on mean depression symptoms, it was not statistically or clinically significant at post-treatment as measured by atypical depression supplement score (p=0.26) or the HRDS (p=0.76,
Table 356) and the quality of evidence was very low due to serious imprecision and risk of bias.

Table 356: Summary of findings tables for treatment effects of massage versus support on depression outcomes

Bright light therapy compared with placebo for depression in pregnancy and the postnatal period

Patient or population: patients with Depression in pregnancy and the postnatal period
Settings:
Intervention: Bright light therapy
Comparison: Placebo

Outcomes	Illustrative comparative risks* (95% CI) Assumed risk Placebo	Corresponding risk Bright light therapy	Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
Depressive symptoms at post-treatment (5 weeks) - SIGH-ADS-29 (atypical depression supplement) score		The mean depressive symptoms at post-treatment (5 weeks) - sigh-ads-29 (atypical depression supplement) score in the intervention groups was 0.45 standard deviations lower (1.23 lower to 0.33 higher)		27 (1 study)	⊕⊖⊖⊖ very low1,2	SMD -0.45 (-1.23 to 0.33)

Depressive symptoms at post-treatment (5 weeks) - HDRS-17 score	The mean depressive symptoms at post-treatment (5 weeks) - hdrs-17 score in the intervention groups was 0.16 standard deviations lower (0.93 lower to 0.6 higher)	27 (1 study)	⊕⊖⊖⊖ very low1,2	SMD -0.16 (-0.93 to 0.6)

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Risk of bias in several domains
2 Optimal information size (for dichotomous outcomes, OIS=300 events; for continuous outcomes, OIS=400 participants) not met.

Anxiety outcomes (by intervention)

Physical activity versus control

There was no statistically or clinically significant effect of [physical activity on mean anxiety scores at post-treatment (p=0.43, Table 357).

Table 357: Summary of findings table for the effects of physical interventions on anxiety in pregnancy and the postnatal period

Anxiety: Physical activity versus control for					
Patient or population: patients with **Settings:** **Intervention: Anxiety: Physical activity versus control**					
Outcomes	**Illustrative comparative risks* (95% CI)** **Assumed risk** **Control**	**Corresponding risk** **Anxiety: Physical activity versus control**	**Relative effect (95% CI)**	**No of Participants (studies)**	**Quality of the evidence (GRADE)**	**Comments**
Anxiety symptoms- Post-treatment (0-9 weeks)- available case analysis		The mean anxiety symptoms- post-treatment (0-9 weeks)- available case analysis in the intervention groups was 0.18 standard deviations higher (0.27 lower to 0.63 higher)		75 (1 study)	⊕⊖⊖⊖ very low1,2	SMD 0.18 (-0.27 to 0.63)

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Unclear risk of bias in several domains
2 95% CI crosses both line of no effect and measure of appreciable benefit or harm (SMD -0.5/0.5) or RR 0.75/1.25 and optimal information size (400 participants) not met

Electroacupuncture versus non-invasive sham acupuncture

There was no statistically or clinically significant effect of electroacupuncture on mean anxiety scores at post-treatment (p=0.96,
Table 358).

Table 358: Summary of findings table for the effects of electroacupuncture on anxiety in pregnancy and the postnatal period

Electroacupuncture compared with non-invasive sham acupuncture for anxiety in pregnancy and the postnatal period

Patient or population: patients with anxiety in pregnancy and the postnatal period
Settings:
Intervention: Electroacupuncture
Comparison: Non-invasive sham acupuncture

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk Non-invasive sham acupuncture	Corresponding risk Electroacupuncture				
Anxiety mean scores - available case analysis		The mean anxiety mean scores - available case analysis in the intervention groups was 0.02 standard deviations lower (0.9 lower to 0.85 higher)		20 (1 study)	⊕⊖⊖⊖ very low1,2	SMD -0.02 (-0.9 to 0.85)

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

1 Risk of bias in several domains
2 Optimal information size (for dichotomous outcomes, OIS=300 events; for continuous outcomes, OIS=400 participants) not met.

General mental health outcomes (by intervention)

Physical activity versus control

There was low quality, single study (N=75) evidence for a statistically significant (p=0.05) beneficial effect of physical activity on mean sleep disturbance score at post-

treatment, however the effect size failed to reach a threshold indicative of clinically significant benefits (
Table 359). In addition the quality of evidence was low due to the serious imprecision (small sample size) and unclear risk of bias in several domains.

Table 359: Summary of findings table for the effects of physical interventions on anxiety in pregnancy and the postnatal period

General mental health: Physical activity versus control for

Patient or population: patients with
Settings:
Intervention: General mental health: Physical activity versus control

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	**Corresponding risk**				
	Control	**General mental health: Physical activity versus control**				
Sleep disturbances- Post-intervention (0-9 weeks)- available case		The mean sleep disturbances- post-intervention (0-9 weeks)- available case in the intervention groups was 0.45 standard deviations lower (0.91 lower to 0.01 higher)		75 (1 study)	⊕⊕⊖⊖ low1,2	SMD -0.45 (-0.91 to 0.01)

*The basis for the assumed risk (for example the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the

estimate of effect and is likely to change the estimate. **Very low quality**: We are very uncertain about the estimate.
1 Unclear risk of bias in several domains 2 Optimal information size (for dichotomous outcomes, OIS=300 events; for continuous outcomes, OIS=400 participants) not met.

8.6.4 Health economics evidence

Systematic literature review

No studies assessing the cost effectiveness of interventions for the treatment of mental health problems in pregnancy or the postnatal period were identified by the systematic search of the economic literature undertaken for this guideline. Details on the methods used for the systematic search of the economic literature are described in Chapter 3.

8.7 ELECTROCONVULSIVE THERAPY FOR MENTAL HEALTH PROBLEMS IN PREGNANCY AND THE POSTNATAL PERIOD

8.7.1 Clinical review protocol (ECT)

The review protocol summary, including the review question(s), information about the databases searched, and the eligibility criteria used for this section of the guideline, can be found in Table 360. A complete list of review questions can be found in Appendix 8; further information about the search strategy can be found in Appendix 10; the full review protocols can be found in Appendix 9.

Table 360: Clinical review protocol summary for the review of ECT

Component	Description
Review question(s)	R.Q. 4.4 For women with mental health problems who are pregnant or in the postnatal period, what are the benefits and/or potential harms of electroconvulsive therapy to treat mental health problems?
Population	Included Women who have mental health problems during pregnancy and the postnatal period (from delivery to the end of the first year). Include: • Women with sub-threshold symptoms • Women with diagnosed mild, moderate and severe disorders Exclude women: • With no current diagnosis of a mental health problem • who are greater than 1 year into the postnatal period • who are not pregnant or postnatal (up to 1 year postnatal)
Intervention(s)	Electroconvulsive therapy
Comparison	Treatment as usual, no treatment, wait-list control, active control, other active interventions

Critical outcomes	Maternal Outcomes Symptom-based Diagnosis of mental health problem Symptomatology Relapse Use of drugs/alcohol Service utilisation Hospitalisation Retention in services (assessed through dropout rates as a proxy measure) Health service utilisation (for instance, use of psychiatric services) Experience of care Satisfaction (validated measures only, specific items will not be analysed) Acceptability of treatment (assessed through questioning or through including dropout as a proxy measure) Quality of life Quality of life measures Functional disability Social functioning Social support Self-esteem Perceived parenting stress Maternal confidence Preservation of rights Harm Side effects (including dropout because of side effects) Maternal mortality and serious morbidity including self-harm and suicide attempts Quality of mother-infant interaction Quality of mother-infant interaction (including maternal sensitivity and child responsivity) Maternal attitude towards motherhood Establishing or continuing breastfeeding Infant outcomes (no restriction on length of follow-up) Fetal and infant physical development (including congenital malformations) Side effects (especially of pharmacological interventions for the fetus and for the infant if breastfeeding) Apgar score Birth weight Admission to neonatal intensive care unit Cognitive development of the infant Emotional development of the infant Physical development of the infant Prevention of neglect or abuse of the infant Optimal care of infant (for example vaccinations, well-baby check-ups) Foetal/infant mortality Foetal/infant morbidity Service use Planned (health visitor, vaccinations, well-baby check-ups)

	Unplanned (emergency department visits, inpatient, urgent or acute care) Social service involvement
Electronic databases	CENTRAL, CINAHL, EMBASE, MEDLINE, PsycINFO
Date searched	Inception to 00.00.2010
Study design	Systematic reviews of RCTs Primary RCTs

8.7.2 Studies considered

No studies assessing the efficacy effectiveness of ECT for women with mental health problems in pregnancy and the postnatal period were identified by the systematic search of the literature undertaken for this guideline.

8.7.3 Health economics evidence

Systematic literature review

No studies assessing the cost effectiveness of ECT for women with mental health problems in pregnancy or the postnatal period were identified by the systematic search of the economic literature undertaken for this guideline. Details on the methods used for the systematic search of the economic literature are described in Chapter 3.

8.8 LINKING EVIDENCE TO RECOMMENDATIONS

8.8.1 Pharmacological interventions for prevention of mental health problems

There was limited and low quality evidence for the prevention of mental health problems in women with no identified risk factors; there was no evidence for a beneficial effect of omega-3, and inconsistent evidence for calcium and selenium on preventing depression. For women with risk factors for depression, there was no evidence for a beneficial effect of thyroxine (for women positive for thyroid antibodies), and a non-beneficial effect of norethisterone (for women with low socio-economic status) on depression outcomes, with evidence for an increased risk of bleeding problems. For the prophylaxis of depression, there was inconsistent evidence for an effect of antidepressants (both SSRIs and TCAs) on preventing recurrence of depression, and some evidence for an increased risk of adverse events associated with both antidepressants, however the quality of evidence was very low. No data were available to the GDG on the cost effectiveness or impact on resource use of the interventions considered in pregnancy. Therefore the GDG judged that no recommendation on prevention of mental health problems in pregnancy and the postnatal period could be made.

8.8.2 Pharmacological interventions for the treatment of mental health problems – harm and efficacy

Antidepressants (TCAs, SSRIs, (S)NRIs)

In reviewing the evidence and developing the recommendations on the harms associated with antidepressant use in pregnancy, the GDG was mindful of the serious nature of the outcomes reviewed, which could have a profound effect on the life course of any individual who is born with a major congenital defect. They were also concerned about the potentially increased rate of a number of the outcomes considered in women with depression who had not been exposed to antidepressant drugs. The GDG were cautious when it came to interpreting the data on individual drugs given the variation in the size of the datasets. Finally, although absolute rate differences were small in most cases, the GDG was aware of the high level of prescribing of antidepressant drugs and the potential impact on a large number of women and babies.

The GDG agreed that there was a small but significant increase in a number of congenital abnormalities (in particular cardiac abnormalities) for a number of important outcomes. However, because of the limitations of the comparator groups (not all contained women with a depressive disorder and where this was the case it was not often known if the severity of the disorder in the comparator group was similar), the GDG was uncertain whether all of this increase could be accounted for by the drug. The GDG considered the possibility that the potential harms arising for the women having the mental health problem may possibly increase as the severity of the depressive disorder increased. The GDG also took into consideration that for many of these women there may be a prior history of depression and this may also be used to guide prescribing practice. Given the considerable uncertainty surrounding the evidence, the GDG adopted a cautious approach in developing its recommendations and also took account of what is known about the effective treatment of depression in non-pregnant women. No data were available to the GDG on the cost effectiveness or impact on resource use of the interventions considered in pregnancy. However, the guideline meta-analysis of clinical evidence points to similar levels of harms across the antidepressants reviewed. Most of the drugs reviewed are off patent and available in generic form. In the case of newer drugs the lack of any greater effect than older drugs makes the added cost potentially not worthwhile. Again the GDG took into account what is known about the cost effectiveness of treatment of depression in non-pregnant women. The GDG also took into account that many women (up to 90%) stop taking medication when they discover that they are pregnant, often without consulting a healthcare professional.

After considering these factors and the significant limitations of the evidence, the GDG decided that for antidepressants (and for most other drugs used for the treatment of mental health problems in pregnancy) that the primary focus of the recommendations should be on a set of principles to guide prescribing rather than a set of recommendations for individual drugs. These principles are as follows:

- All women of childbearing potential should be informed of the limited evidence and consequent uncertainty regarding the harms to the fetus associated with the use of antidepressant medication in pregnancy and the postnatal period, including breastfeeding.
- All women of childbearing potential should be informed of the benefits and side effects associated with the use of antidepressants in pregnancy and the postnatal period (including breastfeeding) if such drugs are being considered.
- All women of childbearing potential should be informed of the background risks associated with depression in pregnancy and the postnatal period.
- All risks should be made clear to women in a manner which is understandable and is based on an assessment of each woman's needs.
- Non-specialists should seek advice or refer onto specialists if they are uncertain about the benefits and harms associated with the use of a particular drug.
- Given the uncertainty about the risks, the threshold for the prescribing of antidepressants should be adjusted in comparison to that for non-pregnant women and that there should be an increased level of monitoring and support for women taking antidepressants in pregnancy and the postnatal period.
- Considerable caution should be exercised when changing or stopping antidepressant drugs in pregnancy and the postnatal period.
- Babies should be monitored for the effects of medication taken in pregnancy and a drug offered that enables the woman to breastfeed if she chooses.
- Specific drugs should only be named where there was evidence to support this.
- That the recommendations for all psychotropic drug use in pregnancy as far as possible should be based on a common set of principles as long as they are supported by the available evidence.

Antipsychotics

In reviewing the evidence and developing the recommendations on the harms associated with antipsychotic use in pregnancy the GDG was, as with antidepressants, mindful of the serious nature of the outcomes reviewed, which could have a profound effect on the life course of any individual who is born with a major congenital defect. They were also aware of the potentially increased rate of a number of the outcomes considered in women with psychosis and bipolar disorder who had not been exposed to antipsychotic drugs. There was some indication from studies of women with a psychotic disorder who were not exposed to medication to support this view. The GDG was also cautious when it came to interpreting the data

on individual drugs given the very limited data available and the variation in the size of the datasets.

The GDG agreed that there was a small but significant increase in a number of congenital abnormalities for a number of important outcomes. However, while this rate was reduced and not significant in a comparator group with the disorder there was still a small increase in the absolute rate and the GDG remained uncertain as to whether the increased rate of abnormality could be accounted for by the drug. For a number of neonatal and obstetric outcomes there was evidence of an increased rate of babies being small for gestational age and increased rates of gestational diabetes, preterm delivery and Caesarean section. Again where data were available for disorder-specific comparisons there was a reduction in the absolute rates of these complications. The GDG was also aware that for many of these women there might be a prior history of psychosis or bipolar disorder and this might also be used to guide prescribing practice. Given the uncertainty surrounding the evidence, the GDG adopted a cautious approach in developing its recommendations and also took account of what is known about the effective treatment of psychosis and bipolar disorder in non-pregnant women. In developing the recommendations the GDG considered, in particular, the potential protective function of antipsychotics in reducing the likelihood of postpartum psychosis. The GDG was of the view that it was particularly important to inform women of the risk of not taking medication in pregnancy if they have a history of severe mental illness. In addition the GDG took into account the evidence elsewhere in this chapter on the risks of harm associated with the use of other drugs (notably lithium and valproate) in developing the recommendations. Given the evidence on gestational diabetes and possible related risk for the fetus, the GDG considered it important that additional and careful monitoring for diabetes and offer of an oral glucose tolerance test should be provided for all pregnant women taking an antipsychotic. No data were available to the GDG on the cost effectiveness or impact on resource use of the interventions considered in pregnancy. The GDG considered the potential resource use implications and high costs associated with the management of congenital abnormalities, neonatal and obstetric complications (that is, babies being small for gestational age and increased rates of gestational diabetes, preterm delivery and Caesarean section); however the GDG was also aware that for many of these women there might be a prior history of psychosis or bipolar disorder and the potential protective function of antipsychotics in reducing the likelihood of costly postpartum psychosis. The GDG found it difficult to judge the net effect to NHS costs. Again the GDG took into account what is known about the cost effectiveness of treatment of antipsychotics in non-pregnant women. Moreover, as with depression, the GDG also took into account that many women stop taking medication when they discover that they are pregnant, often without consulting a healthcare professional.

After considering these factors and the significant limitations of the evidence, the GDG decided that as for antidepressants, the primary focus of the recommendations

for the use of antipsychotics in pregnancy should be on a set of principles to guide prescribing. These are as follows:

- All women of childbearing potential should be informed of the limited evidence and consequent uncertainty regarding the harms to the fetus associated with the use of antipsychotic medication in pregnancy and the postnatal period including breastfeeding.
- All women of childbearing potential should be informed of the benefits and side effects associated with the use of antipsychotics in pregnancy and the postnatal period (including breastfeeding) if such drugs are being considered.
- All women of childbearing potential should be informed of the background risks associated with psychotic disorders in pregnancy and the postnatal period.
- All risks should be made clear to women in a manner which is understandable and is based on an assessment of each woman's needs.
- Non-specialists should seek advice or refer onto specialists if they are uncertain about the benefits and harms associated with the use of a particular drug.
- Given the uncertainty about the risks associated with antipsychotics (for example, gestational diabetes) there should be an increased level of monitoring and support (for example, help with drug-induced weight gain and offer an oral glucose tolerance test) for women taking antipsychotics in pregnancy and the postnatal period.
- Considerable caution should be exercised when changing or stopping antipsychotic drugs in pregnancy and the postnatal period.
- Babies should be monitored for the effects of medication taken in pregnancy and a drug offered that enables the woman to breastfeed if she chooses.
- Specific drugs should only be named where there was evidence to support this.

Anticonvulsants

In reviewing the evidence and developing the recommendations on the harms associated with anticonvulsant use in pregnancy the GDG was mindful of the serious nature of the outcomes reviewed, which could have a profound effect on the life course of any individual who is born with a major congenital defect. They were also aware of the potentially increased rate of a number of the outcomes considered in women with bipolar disorderwho had not been exposed to anticonvulsant drugs. There was some indication from studies reviewed of women with a disorder who were not exposed to medication to support this view. The GDG were aware that the dataset was primarily drawn from women with epilepsy but they did not think that this invalidated the evidence, which was still seen as relevant to women with bipolar disorder. The small number of anticonvulsant drugs used in bipolar disorder and the

relatively large number of studies also meant that the GDG was able to consider the evidence for the three drugs (carbamazepine, valproate and lamotrigine) separately. This is important as there is a clear indication of different patterns of harm associated with each drug. The GDG was of the view that the evidence of significant harms (both congenital and neurodevelopmental) to the fetus associated with valproate was such that it should not be used in the acute or long-term treatment of a mental health problem in women of childbearing potential. There was also evidence of an increased rate of congenital harms associated with carbamazepine albeit not at the same level as valproate and not at a level which would preclude the use of carbamazepine in women of childbearing potential. However, the risk for harm associated with carbamazepine was greater than that observed for lamotrigine and the GDG recommended that carbamazepine is not offered for the treatment of a mental health problem to women who are planning a pregnancy, pregnant or considering breastfeeding. For a woman who is already taking carbamazepine and is planning a pregnancy or becomes pregnant, the GDG recommended that a discussion with the woman should be had about stopping the drug given the possible risk of adverse drug interactions and fetal malformations. The review of lamotrigine did not suggest that there was any significant increase in risk associated with its use in pregnancy. However, the GDG were mindful of the need to check lamotrigine levels frequently during pregnancy and into the postnatal period because they vary substantially at this time.

In considering recommendations concerning valproate, the GDG also took into account a recent audit in an East London psychiatric service (Jones et al., 2014) that shows that standards for prescribing sodium valproate to women of childbearing age, in keeping with NICE guidelines, were not being met. Jones and colleagues (2014) reported that in a trust-wide audit of female inpatients of childbearing age, 13% were prescribed sodium valproate in April 2013, and in a re-audit eight months later there were still 9% of female inpatients of childbearing age prescribed sodium valproate. In addition, and even more concerningly, there was evidence that NICE guidance was not being followed in terms of informing service users about the potential for harm associated with valproate. NICE guidance recommends that where there are no alternatives to sodium valproate, women of childbearing potential must be made aware of the teratogenic properties, have a documented pregnancy test and be prescribed folic acid. However, an audit (and a re-audit eight months later) in this East London NHS Trust revealed that despite recommendations these precautions for prescribing valproate (where there were no alternatives) were not being followed. For instance, 86% (April 2013)/83% (December 2013) of female inpatients of childbearing age prescribed sodium valproate had no documented discussion regarding the teratogenicity and other effects on the developing fetus of sodium valproate exposure, 57% (April 2013)/25% (December 2013) did not have a documented pregnancy test, 86% (April 2013)/92% (December 2013) were not prescribed folic acid, and 86% (April 2013)/58% (December, 2013) had not been provided with contraception. The GDG took the view that assuming that the prescribing practices in this trust were representative of other NHS Trusts (and there

was no reason to think otherwise), and the potential for harm that these results imply, there were grounds to strengthen the valproate recommendations that were made in the previous 2007 guideline. This was also consistent with recommendations made in the bipolar guidance.

In developing the recommendations the GDG considered, in particular, the potential protective function of anticonvulsants in reducing the likelihood of postpartum psychosis. The GDG was of the view that it was particularly important to inform the women of the risk of not taking medication in pregnancy if the women had a history of severe mental illness. In addition the GDG took into account the evidence elsewhere in this chapter on the risks of harm associated with the use of other drugs in developing the recommendations. The GDG considered it important that additional and careful monitoring of plasma drug levels should be undertaken for lamotrigine. No data were available to the GDG on the cost effectiveness or impact on resource use of anticonvulsants considered in pregnancy, however the GDG considered the increased rate of congenital and neurodevelopmental defects associated with valproate (when compared with carbamazepine and lamotrigine) and the potential increase to NHS costs. The GDG also noted that there is an increased rate of congenital harms associated with carbamazepine and potential for substantial resource use. Again the GDG took into account what is known about the cost effectiveness of treatment of anticonvulsants in non-pregnant women. As with other classes of drugs the GDG also took into account that many women stop taking medication when they discover that they are pregnant, often without consulting a healthcare professional.

After considering these factors and the significant limitations of the evidence, the GDG decided that as for antidepressants and antipsychotics, prescribing anticonvulsants should be guided by a set of principles, which are set out in the sections above.

Lithium

Lithium has been used in the treatment of bipolar disorder for over 50 years but the data on harm in pregnancy and the postnatal period is very limited. There was some evidence of a small (7 per 1000) increased risk of congenital abnormalities but it was not possible to obtain a clear picture on increased risk of heart defects despite previous concerns about an association of Ebstein's anomaly with the use of lithium in pregnancy. The GDG therefore felt that lithium could have role in the treatment of bipolar disorder in pregnancy but should only be offered in pregnancy and the postnatal period when an antipsychotic medication has not been effective and healthcare professionals should ensure that the woman knows that lithium levels may be high in breast milk with a risk of toxicity for the baby, lithium levels should be monitored more frequently throughout pregnancy and the postnatal period, and lithium should be stopped during labour and plasma lithium levels checked 12 hours after the woman's last dose.

In developing the recommendations the GDG considered, in particular, the potential protective function of lithium in reducing the likelihood of postpartum psychosis. The GDG was also of the view that it was particularly important to inform the woman of the risk of not taking medication in pregnancy if the woman has a history of severe mental illness. In addition the GDG took into account the evidence elsewhere in this chapter on the risks of harm associated with the use of other drugs in developing the recommendations. No data were available to the GDG on the cost effectiveness or impact on resource use of lithium considered in pregnancy. The GDG found it very difficult to judge the net effect on NHS costs associated with the use of lithium in the treatment of bipolar disorder in pregnancy (that is, the increased risk of congenital abnormalities and the reduced likelihood of postpartum psychosis). Again the GDG took into account what is known about the cost effectiveness of treatment of lithium in non-pregnant women. As with other drugs the GDG also took into account that many women stop taking medication when they discover that they are pregnant, often without consulting a healthcare professional.

After considering these factors and the significant limitations of the evidence, the GDG decided that as for antidepressants and antipsychotics, prescribing lithium should be guided by a set of principles, which are set out in the sections above.

Benzodiazepines and related drugs

Considering the limited evidence for congenital harms and the increase in obstetric complications associated with benzodiazepines and related drugs, the GDG did not consider there to be sufficient evidence of clinical benefit to justify their use in pregnancy and the postnatal period. Furthermore the GDG was of the view, given the potential for harm, that a woman who is taking a benzodiazepine when she becomes pregnant should be encouraged and supported in stopping the medication.

Stimulants

Considering the limited evidence on harms associated with stimulants (only data available for methyphenidate), and its limited use, the GDG did not wish to make any specific recommendations on the use of these drugs.

Treatment options for specific mental health problems

In addition to reviewing the evidence for harms of psychotropic medication, the GDG also reviewed the efficacy of pharmacological interventions in pregnancy and the postnatal period. The evidence for efficacy of psychotropic medication in pregnancy and the postnatal period was limited, both in terms of available studies and in the low quality of the evidence reviewed. The GDG was of the view that the evidence for omega-3 oils and transdermal oestrogen was weak in that there was no clear indication of any benefit. The GDG decided therefore to make no specific recommendations for omega-3 or transdermal oestrogen in the treatment of mental health problems in pregnancy and the postnatal period. The evidence for antidepressants was also limited but broadly in line with the evidence of the efficacy of this medication in non-pregnant populations. The GDG therefore was of the view

that antidepressants had a role to play in the treatment of depression and anxiety disorders in pregnancy and the postnatal period.

The literature review was unable to identify any evidence for the efficacy of antipsychotic medication in pregnancy and the postnatal period. In line with the principles set out below, the GDG therefore referred to existing NICE guidelines. Again, in line with the principles for the reduction of harm, the GDG decided not to single out specific drugs.

The GDG developed an overarching principle regarding interventions to offer or consider for a specific mental health problem. This was arrived at by consensus and states that where a review of the evidence for efficacy of an intervention might be limited, but which contains no indication of a difference in efficacy or harm from the data in non-pregnant populations, then it is reasonable to extrapolate from evidence from non-pregnant populations to inform recommendations for this guideline (see Chapter 3, Section 3.5.6).

The consequence of this is that NICE guidelines for individual mental health problems should be followed other than where specifically indicated in this guideline. In making its recommendations regarding when and how interventions for mental health problems in pregnancy and the postnatal period might need to be modified, the GDG took into account the following evidence:

- reviews undertaken for this guideline update (in this chapter and also in Chapter 7 on psychological and psychosocial interventions)
- their own expert knowledge and opinion
- the recommendations and underlying evidence from the previous 2007 guideline
- NICE guidelines on specific mental health problems, most notably *Depression in Adults* (NICE, 2009a), *Common Mental Health Disorders* (NICE, 2011b), *Bipolar Disorder* (NICE, 2014b) and *Psychosis and Schizophrenia in Adults* (NICE, 2014a), *Drug Misuse: Opioid Detoxification* (NICE, 2007b) and *Alcohol-use Disorders* (NICE, 2011c).

For women with depression in pregnancy or the postnatal period, the GDG judged that for those with a history of severe disorder, but who present with mild depression when pregnant or after childbirth, an antidepressant should be considered as an option, but that healthcare professionals should take into account all of the recommendations regarding balancing risks and benefits. For women who have a moderate to severe episode of depression or an anxiety disorder that has its onset in pregnancy or the postnatal period, the GDG considered that the full range of options recommended in other relevant NICE guidelines should be available, including medication, psychological interventions, and a combination of both. But for women with pre-existing mild to moderate depression, the GDG considered the evidence reviewed for this guideline update in this chapter and in Chapter 7 and

recommended that antidepressant medication should be discontinued and a psychosocial intervention (facilitated self-help) considered. Women with pre-existing mental health problems might be inclined to stop their medication when they know they are pregnant; but for women with moderate to severe depression or an anxiety disorder, the GDG advises changing to medication with lower risk of adverse effects and/or a structured psychological intervention. The GDG wished to emphasise that the clinician will have to carefully balance the need to ensure the woman is offered the optimal treatment against any risks associated with medication or untreated disorder for the fetus.

For women with severe mental illness (psychosis, schizophrenia or bipolar disorder), the GDG judged that an antipsychotic should be offered if a pregnant woman develops mania or psychosis and is not taking any psychotropic medication; the choice of antipsychotic will depend on full consideration of the risks and benefits. For women with pre-existing bipolar disorder, the GDG judged that an antipsychotic in line with the NICE guidelines on bipolar disorder (NICE, 2014b) might be a suitable drug to offer or continue with if a woman plans to breastfeed. Antipsychotic medication should also be offered if a woman with bipolar disorder if stopping the prophylactic use of lithium. If a pregnant woman develops mania while taking prophylactic medication, the GDG considered a number of options, including checking and, if necessary increasing the dose, of the existing medication, changing to antipsychotic medication, lithium if the mania is severe and there has been no response to other medication, and, finally, ECT if there is no response to lithium.

No studies were found that matched the inclusion criteria for the updated review of rapid tranquilisation. Therefore the recommendation from the previous 2007 guideline remains unchanged.

In making recommendations for pregnant women dependent on drugs and alcohol in the light of lack of evidence, the GDG drew on discussion with experts for this important area. They recommend that detoxification for pregnant women carried out in conjunction with specialist mental health and alcohol or substance misuse services, but highlight that women who do not wish to undertake a detoxification should be offered interventions to reduce their drug and alcohol intake. The GDG considered it important that healthcare professionals are aware of the possibility of accidental overdose in women who stop or reduce drug misuse in pregnancy but start misusing again soon after childbirth.

There was also a lack of high quality evidence for pharmacological and psychosocial interventions for sleep problems and insomnia in pregnancy and the postnatal period. However, the GDG was mindful that the previous 2007 guideline recommended low-dose chlorpromazine or amitriptyline for women with 'serious and chronic problems' and wished to amend this. The GDG considered the risks associated with low-dose chlorpromazine or amitriptyline and sedating drugs such as zopiclone, as well as the review of harms associated with both antidepressants

and antipsychotics, and judged that promethazine was a suitable alternative in pregnancy.

8.8.3 Physical interventions

The evidence for physical interventions was limited and the quality of evidence low. In reviewing the available data there appeared to be some beneficial effects of physical activity on preventing depression. There was limited evidence for a beneficial effect of physical activity for the treatment of depression however there was some evidence for depression-specific acupuncture and bright-light therapy. However the GDG did not feel the evidence was strong enough to make any specific recommendations about physical interventions.

No studies were found that matched the inclusion criteria for the updated review of ECT. Therefore the recommendation from the previous 2007 guideline remains unchanged, other than to use current NICE style for recommendations. The summary from the previous guideline stated that: 'The use of ECT during pregnancy is not well researched, although some complications for mother and fetus have been described, including transient, self-limited disturbances in fetal cardiac rhythm, suspected vaginal bleeding, uterine contractions (although these did not result in premature labour or adverse consequences, severe abdominal pain directly after ECT treatments was reported in pregnant women – though the babies were born healthy) and premature labour (Miller, 1995). Five cases of congenital anomalies in offspring prenatally exposed to ECT have been reported, including hypertelorism, optic atrophy, anencephaly, clubbed foot and pulmonary cysts, although these were not considered the direct result of ECT (Miller, 1995). The risks of ECT therefore need to be balanced against the risks of using alternative treatments, in consultation with anaesthetist and obstetrician. ECT was cautiously recommended in the NICE Technology Appraisal (NICE, 2003).

8.9 RECOMMENDATIONS

8.9.1 Clinical recommendations

Using and modifying NICE guidelines for specific mental health problems

Assessment and treatment in pregnancy and the postnatal period

8.9.1.1 Use this guideline in conjunction with the NICE guideline for a specific mental health problem (see the related NICE guidance in section 3.2 [in the NICE guideline]) to inform assessment and treatment decisions in pregnancy and the postnatal period, and take into account:

- any variations in the nature and presentation of the mental health problem in pregnancy or the postnatal period
- the setting for assessment and treatment (for example, primary or secondary care services or in the community, the home or remotely by phone or computer)
- recommendations 5.4.8.5 5.4.8.5 to5.4.8.10 in this guideline on assessment in pregnancy and the postnatal period
- recommendations 8.9.1.6 to 8.9.1.33in this guideline on starting, using and stopping treatment in pregnancy and the postnatal period
- recommendations 5.4.8.13 to 5.4.8.14, 7.7.1.6 to 7.7.1.17 and 8.9.1.35 to 8.9.1.48 in this guideline on treating specific mental health problems in pregnancy and the postnatal period. **[new 2014]**

Considerations for women of childbearing potential

8.9.1.2 When prescribing psychotropic medication for women of childbearing potential, take account of the latest data on the risks to the fetus and baby. **[new 2014]**

8.9.1.3 Do not offer valproate for acute or long-term treatment of a mental health problem in women of childbearing potential. **[new 2014]**

Treatment decisions, advice and monitoring for women who are planning a pregnancy, pregnant or in the postnatal period

Information and advice

8.9.1.4 Consider referring a woman to a secondary mental health service (preferably a specialist perinatal mental health service) for preconception counselling if she has a current or past severe mental health problem and is planning a pregnancy. **[new 2014]**

8.9.1.5 Discuss breastfeeding with all women who may need to take psychotropic medication in pregnancy or in the postnatal period. Explain to them the benefits of breastfeeding, the potential risks associated with taking psychotropic medication when breastfeeding and with stopping some medications in order to breastfeed. Discuss treatment options that would enable a woman to breastfeed if she wishes and support women who choose not to breastfeed. **[new 2014]**

Starting, using and stopping treatment

General advice

8.9.1.6 Before starting any treatment in pregnancy and the postnatal period, discuss with the woman the higher threshold for pharmacological interventions arising from the changing risk–benefit ratio for psychotropic medication at this time and the likely benefits of a psychological intervention. **[new 2014]**

8.9.1.7 If the optimal treatment for a woman with a mental health problem is psychotropic medication combined with a psychological intervention, but she declines or stops taking psychotropic medication in pregnancy or the postnatal period, ensure that:

- she is adequately supported and
- has the opportunity to discuss the risk associated with stopping psychotropic medication and
- is offered, or can continue with, a psychological intervention. **[new 2014]**

8.9.1.8 When psychotropic medication is started in pregnancy and the postnatal period, consider seeking advice, preferably from a specialist in perinatal mental health, and:

- choose the drug with the lowest risk profile for the woman, fetus and baby, taking into account a woman's previous response to medication
- use the lowest effective dose (this is particularly important when the risks of adverse effects to the woman, fetus and baby may be dose related), but note that sub-therapeutic doses may also expose the fetus to risks and not treat the mental health problem effectively
- use a single drug, if possible, in preference to 2 or more drugs
- take into account that dosages may need to be adjusted in pregnancy.**[2014]**

8.9.1.9 When a woman with severe mental illness decides to stop psychotropic medication in pregnancy and the postnatal period, discuss with her:

- her reasons for doing so
- the possibility of:

- restarting the medication
- switching to other medication
- having a psychological intervention

• increasing the level of monitoring and support.

Ensure she knows about any risks to herself, the fetus or baby when stopping medication. [**new 2014**]

8.9.1.10 When a woman with depression or an anxiety disorder decides to stop taking psychotropic medication in pregnancy and the postnatal period, discuss with her:

• her reasons for doing so

• the possibility of:

- having a psychological intervention
- restarting the medication if the depression or anxiety disorder is or has been severe and there has been a previous good response to treatment
- switching to other medication

• increasing the level of monitoring and support while she is not taking any medication.

Ensure she knows about any risks to herself, the fetus or baby when stopping medication. **[new 2014]**

8.9.1.11 If a pregnant woman has taken psychotropic medication with known teratogenic risk at any time in the first trimester:

- confirm the pregnancy as soon as possible
- explain that stopping or switching the medication after pregnancy is confirmed may not remove the risk of fetal malformations
- offer screening for fetal abnormalities and counselling about continuing the pregnancy
- explain the need for additional monitoring and the risks to the fetus if she continues to take the medication.

Seek advice from a specialist if there is uncertainty about the risks associated with specific drugs. **[new 2014]**

TCAs, SSRIs, (S)NRIs

8.9.1.12 When choosing a tricyclic antidepressant (TCA), selective serotonin reuptake inhibitor (SSRI) or (serotonin-) noradrenaline reuptake inhibitor [(S)NRI][20], take into account:

---

[20] Although this use is common in UK clinical practice, at the time of publication (December 2014]), TCAs, SSRIs and (S)NRIs did not have a UK marketing authorisation for this indication. The prescriber should follow relevant professional guidance, taking full responsibility for the decision. Informed consent should be obtained and documented. See the General Medical Council's Good practice in prescribing and managing medicines and devices for further information.

- the woman's previous response to these drugs
- the stage of pregnancy
- what is known about their reproductive safety (for example, fetal cardiac abnormalities and persistent pulmonary hypertension in the newborn baby)
- the uncertainty about whether any increased risk to the fetus and other problems for the woman or baby can be attributed directly to these drugs or may be caused by other factors
- the risk of discontinuation symptoms in the woman and neonatal adaptation syndrome in the baby with most TCAs, SSRIs and (S)NRIs, in particular paroxetine and venlafaxine**. [new 2014]**

8.9.1.13 When assessing the risks and benefits of TCAs, SSRIs or (S)NRIs[21] for a woman who is considering breastfeeding, take into account:

- the benefits of breastfeeding for the woman and baby
- the uncertainty about the safety of these drugs for the breastfeeding baby
- the risks associated with switching from or stopping a previously effective medication.

Seek advice from a specialist (preferably from a specialist perinatal mental health service) if there is uncertainty about specific drugs. **[new 2014]**

Benzodiazepines

8.9.1.14 Do not offer benzodiazepines to women in pregnancy and the postnatal period except for the short-term treatment of severe anxiety and agitation. **[2014]**

8.9.1.15 Consider gradually stopping benzodiazepines in women who are planning a pregnancy, pregnant or considering breastfeeding. **[2014]**

[21] Although this use is common in UK clinical practice, at the time of publication (December 2014]), TCAs, SSRIs and (S)NRIs did not have a UK marketing authorisation for this indication. The prescriber should follow relevant professional guidance, taking full responsibility for the decision. Informed consent should be obtained and documented. See the General Medical Council's Good practice in prescribing and managing medicines and devices for further information.

Antipsychotic medication

8.9.1.16 When assessing the risks and benefits of antipsychotic medication[22] for a pregnant woman, take into account risk factors for gestational diabetes and excessive weight gain**. [new 2014]**

8.9.1.17 When choosing an antipsychotic, take into account that there are limited data on the safety of these drugs in pregnancy and the postnatal period. **[new 2014]**

8.9.1.18 Measure prolactin levels in women who are taking prolactin-raising antipsychotic medication and planning a pregnancy, because raised prolactin levels reduce the chances of conception. If prolactin levels are raised, consider a prolactin-sparing antipsychotic. **[2014]**

8.9.1.19 If a pregnant woman is stable on an antipsychotic and likely to relapse without medication, advise her to continue the antipsychotic. **[new 2014]**

8.9.1.20 Advise pregnant women taking antipsychotic medication about diet and monitor for excessive weight gain, in line with the guideline on weight management before, during and after pregnancy (NICE public health guidance 27). **[new 2014]**

8.9.1.21 Monitor for gestational diabetes in pregnant women taking antipsychotic medication in line with the guideline on diabetes in pregnancy (NICE clinical guideline 63) and offer an oral glucose tolerance test. **[new 2014]**

8.9.1.22 Do not offer depot antipsychotics to a woman who is planning a pregnancy, pregnant or considering breastfeeding, unless she is responding well to a depot and has a previous history of non-adherence with oral medication. **[new 2014]**

Anticonvulsants (valproate, carbamazepine and lamotrigine)

8.9.1.23 Do not offer valproate for acute or long-term treatment of a mental health problem in women who are planning a pregnancy, pregnant or considering breastfeeding. **[new 2014]**

8.9.1.24 If a woman is already taking valproate and is planning a pregnancy, advise her to gradually stop the drug because of the risk of fetal malformations and adverse neurodevelopment outcomes after any exposure in pregnancy. **[2014]**

---

[22] Although this use is common in UK clinical practice, at the time of publication (December 2014]), antipsychotic medication did not have a UK marketing authorisation for this indication. The prescriber should follow relevant professional guidance, taking full responsibility for the decision. Informed consent should be obtained and documented. See the General Medical Council's Good practice in prescribing and managing medicines and devices for further information.

8.9.1.25 If a woman is already taking valproate and becomes pregnant, stop the drug because of the risk of fetal malformations and adverse neurodevelopmental outcomes. **[2014]**

8.9.1.26 Do not offer carbamazepine to treat a mental health problem in women who are planning a pregnancy, pregnant or considering breastfeeding. **[new 2014]**

8.9.1.27 If a woman is already taking carbamazepine and is planning a pregnancy or becomes pregnant, discuss with the woman the possibility of stopping the drug (because of the risk of adverse drug interactions and fetal malformations). **[new 2014]**

8.9.1.28 If a woman is taking lamotrigine[23] during pregnancy, check lamotrigine levels frequently during pregnancy and into the postnatal period because they vary substantially at these times. **[new 2014]**

Lithium

8.9.1.29 Do not offer lithium[24] to women who are planning a pregnancy or pregnant, unless antipsychotic medication has not been effective. **[new 2014]**

8.9.1.30 If antipsychotic medication has not been effective and lithium is offered to a woman who is planning a pregnancy or pregnant, ensure:

- the woman knows that there is a risk of fetal heart malformations when lithium is taken in the first trimester, but the size of the risk is uncertain
- the woman knows that lithium levels may be high in breast milk with a risk of toxicity for the baby
- lithium levels are monitored more frequently throughout pregnancy and the postnatal period. **[new 2014]**

8.9.1.31 If a woman taking lithium becomes pregnant, consider stopping the drug gradually over 4 weeks if she is well. Explain to her that:

- stopping medication may not remove the risk of fetal heart malformations

---

[23]At the time of publication (December 2014), lamotrigine did not have a UK marketing authorisation for this indication. The prescriber should follow relevant professional guidance, taking full responsibility for the decision. Informed consent should be obtained and documented. See the General Medical Council's Good practice in prescribing and managing medicines and devicesfor further information.

[24] Although this use is common in UK clinical practice, at the time of publication (December 2014), lithium did not have a UK marketing authorisation for this indication. The prescriber should follow relevant professional guidance, taking full responsibility for the decision. Informed consent should be obtained and documented. See the General Medical Council's Good practice in prescribing and managing medicines and devices for further information.

- there is a risk of relapse, particularly in the postnatal period, if she has bipolar disorder. **[2014]**

8.9.1.32 If a woman taking lithium becomes pregnant and is not well or is at high risk of relapse, consider:

- switching gradually to an antipsychotic or
- stopping lithium and restarting it in the second trimester (if the woman is not planning to breastfeed and her symptoms have responded better to lithium than to other drugs in the past) or
- continuing with lithium if she is at high risk of relapse and an antipsychotic is unlikely to be effective. **[new 2014]**

8.9.1.33 If a woman continues taking lithium during pregnancy:

- check plasma lithium levels every 4 weeks, then weekly from the 36th week
- adjust the dose to keep plasma lithium levels in the woman's therapeutic range
- ensure the woman maintains an adequate fluid balance
- ensure the woman gives birth in hospital
- ensure monitoring by the obstetric team when labour starts, including checking plasma lithium levels and fluid balance because of the risk of dehydration and lithium toxicity
- stop lithium during labour and check plasma lithium levels 12 hours after her last dose. **[2014]**

Providing assessment and interventions for mental health problems in pregnancy and the postnatal period

8.9.1.34 When offering psychotropic medication during pregnancy and the postnatal period, follow the principles in recommendations 8.9.1.6 to 8.9.1.33. **[new 2014]**

Treating specific mental health problems in pregnancy and the postnatal period

Interventions for depression

8.9.1.35 For a woman with a history of severe depression who initially presents with mild depression in pregnancy or the postnatal period, consider a TCA, SSRI or (S)NRI. **[new 2014]**

8.9.1.36 For a woman with moderate or severe depression in pregnancy or the postnatal period, consider the following options:

- a high-intensity psychological intervention (for example, CBT)
- a TCA, SSRI or (S)NRI if the woman understands the risks associated with the medication and the mental health problem in pregnancy and the postnatal period and:

  - she has expressed a preference for medication or
  - she declines psychological interventions or
  - her symptoms have not responded to psychological interventions
- a high-intensity psychological intervention in combination with medication if the woman understands the risks associated with the medication and the mental health problem in pregnancy and the postnatal period and there is no response, or a limited response, to a high-intensity psychological intervention or medication alone. **[new 2014]**

8.9.1.37 If a woman who is taking a TCA, SSRI or (S)NRI for mild to moderate depression becomes pregnant, discuss stopping the medication gradually and consider facilitated self-help (delivered as described in recommendation 1.4.2.2 of the guideline on depression in adults [NICE clinical guideline 90]). **[new 2014]**

8.9.1.38 If a pregnant woman is taking a TCA, SSRI or (S)NRI for moderate depression and wants to stop her medication, take into account previous response to treatment, stage of pregnancy, risk of relapse, risk associated with medication and her preference, and discuss with her the following options:

- switching to a high-intensity psychological intervention (for example, CBT)
- changing medication if there is a drug that is effective for her with a lower risk of adverse effects. [new 2014]

8.9.1.39 If a pregnant woman is taking a TCA, SSRI or (S)NRI for severe depression, take into account previous response to treatment, stage of pregnancy, risk of relapse and risk associated with medication and her preference, and discuss with her the following options:

- continuing with the current medication
- changing medication if there is a drug that is effective for her with a lower risk of adverse effects
- combining medication with a high-intensity psychological intervention (for example, CBT)
- switching to a high-intensity psychological intervention (for example, CBT) if she decides to stop taking medication. **[new 2014]**

Interventions for anxiety disorders

8.9.1.40 If a woman who is taking a TCA, SSRI or (S)NRI for an anxiety disorder becomes pregnant, discuss with her the following options:

- stopping the medication gradually and switching to a high-intensity psychological intervention (for example, CBT)
- continuing with medication if she understands the risks associated with the medication and the mental health problem in pregnancy and the postnatal period and:
  - has expressed a preference for medication or
  - declines psychological interventions or
  - her symptoms have not responded to psychological interventions
- changing medication if there is a drug that is effective for her with a lower risk of adverse effects
- combining medication with a high-intensity psychological intervention (for example, CBT) if the woman understands the risks associated with the medication and the mental health problem in pregnancy and the postnatal period and there is no response, or a limited response, to a high-intensity psychological intervention alone. **[new 2014]**

Interventions for alcohol and drug misuse

8.9.1.41 Offer assisted alcohol withdrawal in collaboration with specialist mental health and alcohol services (preferably in an inpatient setting) to pregnant women who are dependent on alcohol. Work with a woman who does not want assisted alcohol withdrawal to help her reduce her alcohol intake. **[new 2014]**

8.9.1.42 Offer detoxification in collaboration with specialist mental health and substance misuse services to pregnant women who are dependent on opioids. Monitor closely after completion of detoxification. Work with a woman who does not want detoxification to help her reduce her opioid intake. Recognise the risk of accidental overdose in women who stop or reduce drug misuse in pregnancy but start misusing again after childbirth. **[new 2014]**

Interventions for severe mental illness

8.9.1.43 If a pregnant woman develops mania or psychosis and is not taking psychotropic medication, offer an antipsychotic. [**new 2014**]

8.9.1.44 Offer antipsychotic medication in line with recommendations 1.5.3 and 1.5.4 of the guideline on bipolar disorder (NICE clinical guideline 185) as prophylactic medication if a woman with bipolar disorder:

- becomes pregnant and is stopping lithium, or
- plans to breastfeed. **[new 2014]**

8.9.1.45 If a pregnant woman with bipolar disorder develops mania while taking prophylactic medication:

- check the dose of the prophylactic medication and adherence
- increase the dose if the prophylactic medication is an antipsychotic
- suggest changing to an antipsychotic if she is taking another type of prophylactic medication
- consider lithium if there is no response to an increase in dose or change of drug and the woman has severe mania
- consider electroconvulsive therapy (ECT) if there has been no response to lithium. **[new 2014]**

Interventions for sleep problems

8.9.1.46 Advise pregnant women who have a sleep problem about sleep hygiene (including having a healthy bedtime routine, avoiding caffeine and reducing activity before sleep). For women with a severe or chronic sleep problem, consider promethazine[25]. **[new 2014]**

Electroconvulsive therapy

8.9.1.47 Consider electroconvulsive therapy (ECT) for pregnant women with severe depression, severe mixed affective states or mania, or catatonia, whose physical health or that of the fetus is at serious risk. **[2014]**

Rapid tranquillisation

8.9.1.48 A pregnant woman requiring rapid tranquillisation should be treated according to the NICE clinical guidelines on the short-term management of disturbed/violent behaviour, schizophrenia and bipolar disorder (see the related NICE guidance in section 3.2 for details [in the NICE guideline]), except that:

- she should not be secluded after rapid tranquillisation
- restraint procedures should be adapted to avoid possible harm to the fetus
- when choosing an agent for rapid tranquillisation in a pregnant woman, an antipsychotic or a benzodiazepine with a short half-life should be considered; if an antipsychotic is used, it should be at the minimum effective dose because of neonatal extrapyramidal symptoms; if a benzodiazepine is used, the risks of floppy baby syndrome should be taken into account
- during the perinatal period, the woman's care should be managed in close collaboration with a paediatrician and an anaesthetist. **[2007]**

Considerations for women and their babies in the postnatal period

Reviewing treatment for women with severe mental illness

8.9.1.49 After childbirth, review and assess the need for starting, restarting or adjusting psychotropic medication as soon as a woman with a past or present severe mental illness is medically stable. [**new 2014]**

Monitoring babies for effects of psychotropic medication taken in pregnancy

8.9.1.50 If a woman has taken psychotropic medication during pregnancy, carry out a full neonatal assessment of the newborn baby, bearing in mind:

- the variation in the onset of adverse effects of psychotropic medication
- the need for further monitoring
- the need to inform relevant healthcare professionals and the woman and her partner, family or carer of any further monitoring, particularly if the woman has been discharged early**. [new 2014]**

Care of women and their babies if there has been alcohol or drug misuse in pregnancy

8.9.1.51 If there has been alcohol or drug misuse in pregnancy, offer treatment and support after childbirth to both the woman and the baby, including:

- a full neonatal assessment for any congenital abnormalities or neonatal adaptation syndrome
- continuing psychological treatment and support for the woman
- monitoring of the baby. **[new 2014]**

Psychotropic medication and breastfeeding

8.9.1.52 Encourage women with a mental health problem to breastfeed, unless they are taking carbamazepine, clozapine or lithium (valproate is not recommended to treat a mental health problem in women of childbearing potential). However, support each woman in the choice of feeding method that best suits her and her family. **[new 2014]**

8.9.1.53 When assessing the risks and benefits of TCAs, SSRIs or (S)NRIs for women who are breastfeeding, take into account:

- the limited data about the safety of these drugs
- the risks associated with switching from a previously effective medication.

Seek advice from a specialist (preferably from a specialist perinatal mental health service) if needed for specific drugs. **[new 2014]**

8.9.1.54 When assessing the risks and benefits of antipsychotic medication for women who are breastfeeding, take into account:

- the limited data on the safety of these drugs and
- the level of antipsychotic medication in breast milk depends on the drug. **[new 2014]**

8.9.1.55 If a woman is taking psychotropic medication while breastfeeding, monitor the baby for adverse effects. **[2014]**

8.9.2 Research recommendation

8.9.2.1 How safe are drugs used to treat bipolar disorder in pregnancy and the postnatal period?

9 SUMMARY OF RECOMMENDATIONS

9.1 USING THIS GUIDELINE IN CONJUNCTION WITH OTHER NICE GUIDELINES

9.1.1 Improving the experience of care

9.1.1.1 Use this guideline in conjunction with the guidance on service user experience in adult mental health (NICE clinical guideline 136) and patient experience in adult NHS services (NICE clinical guideline 138) to improve the experience of care for women with a mental health problem in pregnancy or the postnatal period. **[new 2014]**

9.1.2 Assessment and treatment in pregnancy and the postnatal period

9.1.2.1 Use this guideline in conjunction with the NICE guideline for a specific mental health problem (see the related NICE guidance in section 3.2 [in the NICE guideline]) to inform assessment and treatment decisions in pregnancy and the postnatal period, and take into account:

- any variations in the nature and presentation of the mental health problem in pregnancy or the postnatal period
- the setting for assessment and treatment (for example, primary or secondary care services or in the community, the home or remotely by phone or computer)
- recommendations 9.6.1.1 to 9.6.1.6 in this guideline on assessment in pregnancy and the postnatal period
- recommendations 9.4.2.1 to 9.4.2.28 in this guideline on starting, using and stopping treatment in pregnancy and the postnatal period
- recommendations 7.7.1.6 to 9.8.8.1 in this guideline on treating specific mental health problems in pregnancy and the postnatal period. **[new 2014]**

9.2 CONSIDERATIONS FOR WOMEN OF CHILDBEARING POTENTIAL

9.2.1.1 Discuss with all women of childbearing potential who have a new, existing or past mental health problem:

- the use of contraception and any plans for a pregnancy

- how pregnancy and childbirth might affect a mental health problem, including the risk of relapse
- how a mental health problem and its treatment might affect the woman, the fetus and baby
- how a mental health problem and its treatment might affect parenting. **[new 2014]**

9.2.1.2 When prescribing psychotropic medication for women of childbearing potential, take account of the latest data on the risks to the fetus and baby. **[new 2014]**

9.2.1.3 Do not offer valproate for acute or long-term treatment of a mental health problem in women of childbearing potential. **[new 2014]**

9.3 PRINCIPLES OF CARE IN PREGNANCY AND THE POSTNATAL PERIOD

9.3.1 Supporting women and their partners, families and carers

9.3.1.1 Acknowledge the woman's role in caring for her baby and support her to do this in a non-judgmental and compassionate way. **[new 2014]**

9.3.1.2 Involve the woman and, if she agrees, her partner, family or carer, in all decisions about her care and the care of her baby. **[new 2014]**

9.3.1.3 When working with girls and young women with a mental health problem in pregnancy or the postnatal period:

- be familiar with local and national guidelines on confidentiality and the rights of the child
- be aware of the recommendations in section 1.4 of the guideline on pregnancy and complex social factors (NICE clinical guideline 110)
- ensure continuity of care for the mental health problem if care is transferred from adolescent to adult services. **[new 2014]**

9.3.1.4 Take into account and, if appropriate, assess and address the needs of partners, families and carers that might affect a woman with a mental health problem in pregnancy and the postnatal period. These include:

- the welfare of the baby and other dependent children and adults
- the role of the partner, family or carer in providing support
- the potential effect of any mental health problem on the woman's relationship with her partner, family or carer. **[new 2014]**

9.3.2 Coordinated care

9.3.2.1 Develop an integrated care plan for a woman with a mental health problem in pregnancy and the postnatal period that sets out:

- the care and treatment for the mental health problem
- the roles of all healthcare professionals, including who is responsible for:

- coordinating the integrated care plan
- the schedule of monitoring
- providing the interventions and agreeing the outcomes with the woman. **[new 2014]**

9.3.2.2 The healthcare professional responsible for coordinating the integrated care plan should ensure that:

- everyone involved in a woman's care is aware of their responsibilities
- there is effective sharing of information with all services involved and with the woman herself
- mental health (including mental wellbeing) is taken into account as part of all care plans
- all interventions for mental health problems are delivered in a timely manner, taking into account the stage of the pregnancy or age of the baby. **[new 2014]**

9.4 TREATMENT DECISIONS, ADVICE AND MONITORING FOR WOMEN WHO ARE PLANNING A PREGNANCY, PREGNANT OR IN THE POSTNATAL PERIOD

9.4.1 Information and advice

9.4.1.1 Provide culturally relevant information on mental health problems in pregnancy and the postnatal period to the woman and, if she agrees, her partner, family or carer. Ensure that the woman understands that mental health problems are not uncommon during these periods and instil hope about treatment. **[new 2014]**

9.4.1.2 Consider referring a woman to a secondary mental health service (preferably a specialist perinatal mental health service) for preconception counselling if she has a current or past severe mental health problem and is planning a pregnancy. **[new 2014]**

9.4.1.3 Discuss treatment and prevention options and any particular concerns the woman has about the pregnancy or the fetus or baby. Provide information to the woman and, if she agrees, her partner, family or carer, about:

- the potential benefits of psychological interventions and psychotropic medication

- the possible consequences of no treatment
- the possible harms associated with treatment
- what might happen if treatment is changed or stopped, particularly if psychotropic medication is stopped abruptly. **[new 2014]**

9.4.1.4 Discuss breastfeeding with all women who may need to take psychotropic medication in pregnancy or in the postnatal period. Explain to them the benefits of breastfeeding, the potential risks associated with taking psychotropic medication when breastfeeding and with stopping some medications in order to breastfeed. Discuss treatment options that would enable a woman to breastfeed if she wishes and support women who choose not to breastfeed. **[new 2014]**

9.4.1.5 If needed, seek more detailed advice about the possible risks of mental health problems or the benefits and harms of treatment in pregnancy and the postnatal period from a secondary mental health service (preferably a specialist perinatal mental health service). This might include advice on the risks and possible harms of taking psychotropic medication while breastfeeding and how medication might affect a woman's ability to care for her baby (for example, sedation). **[new 2014]**

9.4.1.6 Mental health professionals providing detailed advice about the possible risks of mental health problems or the benefits and harms of treatment in pregnancy and the postnatal period should include discussion of the following, depending on individual circumstances:

- the uncertainty about the benefits, risks and harms of treatments for mental health problems in pregnancy and the postnatal period
- the likely benefits of each treatment, taking into account the severity of the mental health problem
- the woman's response to any previous treatment
- the background risk of harm to the woman and the fetus or baby associated with the mental health problem and the risk to mental health and parenting associated with no treatment
- the possibility of the sudden onset of symptoms of mental health problems in pregnancy and the postnatal period, particularly in the first few weeks after childbirth (for example, in bipolar disorder)
- the risks or harms to the woman and the fetus or baby associated with each treatment option
- the need for prompt treatment because of the potential effect of an untreated mental health problem on the fetus or baby
- the risk or harms to the woman and the fetus or baby associated with stopping or changing a treatment. **[new 2014]**

9.4.1.7 When discussing likely benefits and risks of treatment with the woman and, if she agrees, her partner, family or carer:

- acknowledge the woman's central role in reaching a decision about her treatment and that the role of the professional is to inform that decision with balanced and up-to-date information and advice
- use absolute values based on a common denominator (that is, numbers out of 100 or 1000)
- acknowledge and describe, if possible, the uncertainty around any estimate of risk, harm or benefit
- use high-quality decision aids in a variety of numerical and pictorial formats that focus on a personalised view of the risks and benefits, in line with the guidance on patient experience in adult NHS services (NICE clinical guideline 138)
- consider providing records of the consultation, in a variety of visual, verbal or audio formats. **[new 2014]**

9.4.1.8 Healthcare professionals working in universal services and those caring for women in mental health services should:

- assess the level of contact and support needed by women with a mental health problem (current or past) and those at risk of developing one
- agree the level of contact and support with each woman, including those who are not having treatment for a mental health problem
- monitor regularly for symptoms throughout pregnancy and the postnatal period, particularly in the first few weeks after childbirth. **[new 2014]**

9.4.1.9 Discuss and plan how symptoms will be monitored (for example, by using validated self-report questionnaires, such as the Edinburgh Postnatal Depression Scale [EPDS], Patient Health Questionnaire [PHQ-9] or the 7-item Generalized Anxiety Disorder scale [GAD-7]). **[new 2014]**

9.4.2 Starting, using and stopping treatment

General advice

9.4.2.1 Before starting any treatment in pregnancy and the postnatal period, discuss with the woman the higher threshold for pharmacological interventions arising from the changing risk–benefit ratio for psychotropic medication at this time and the likely benefits of a psychological intervention. **[new 2014]**

9.4.2.2 If the optimal treatment for a woman with a mental health problem is psychotropic medication combined with a psychological intervention, but she declines or stops taking psychotropic medication in pregnancy or the postnatal period, ensure that:

- she is adequately supported and

- has the opportunity to discuss the risk associated with stopping psychotropic medication and
- is offered, or can continue with, a psychological intervention. **[new 2014]**

9.4.2.3 When psychotropic medication is started in pregnancy and the postnatal period, consider seeking advice, preferably from a specialist in perinatal mental health, and:

- choose the drug with the lowest risk profile for the woman, fetus and baby, taking into account a woman's previous response to medication
- use the lowest effective dose (this is particularly important when the risks of adverse effects to the woman, fetus and baby may be dose related), but note that sub-therapeutic doses may also expose the fetus to risks and not treat the mental health problem effectively
- use a single drug, if possible, in preference to 2 or more drugs
- take into account that dosages may need to be adjusted in pregnancy. **[2014]**

9.4.2.4 When a woman with severe mental illness decides to stop psychotropic medication in pregnancy and the postnatal period, discuss with her:

- her reasons for doing so
- the possibility of:
  - restarting the medication
  - switching to other medication
  - having a psychological intervention
- increasing the level of monitoring and support.

Ensure she knows about any risks to herself, the fetus or baby when stopping medication. **[new 2014]**

9.4.2.5 When a woman with depression or an anxiety disorder decides to stop taking psychotropic medication in pregnancy and the postnatal period, discuss with her:

- her reasons for doing so
- the possibility of:
  - having a psychological intervention
  - restarting the medication if the depression or anxiety disorder is or has been severe and there has been a previous good response to treatment
  - switching to other medication

- increasing the level of monitoring and support while she is not taking any medication.

Ensure she knows about any risks to herself, the fetus or baby when stopping medication. **[new 2014]**

9.4.2.6 If a pregnant woman has taken psychotropic medication with known teratogenic risk at any time in the first trimester:

- confirm the pregnancy as soon as possible
- explain that stopping or switching the medication after pregnancy is confirmed may not remove the risk of fetal malformations
- offer screening for fetal abnormalities and counselling about continuing the pregnancy
- explain the need for additional monitoring and the risks to the fetus if she continues to take the medication.

Seek advice from a specialist if there is uncertainty about the risks associated with specific drugs. **[new 2014]**

TCAs, SSRIs, (S)NRIs

9.4.2.7 When choosing a tricyclic antidepressant (TCA), selective serotonin reuptake inhibitor (SSRI) or (serotonin-) noradrenaline reuptake inhibitor [(S)NRI][26], take into account:

- the woman's previous response to these drugs
- the stage of pregnancy
- what is known about the reproductive safety of these drugs (for example, the risk of fetal cardiac abnormalities and persistent pulmonary hypertension in the newborn baby)
- the uncertainty about whether any increased risk to the fetus and other problems for the woman or baby can be attributed directly to these drugs or may be caused by other factors
- the risk of discontinuation symptoms in the woman and neonatal adaptation syndrome in the baby with most TCAs, SSRIs and (S)NRIs, in particular paroxetine and venlafaxine. **[new 2014]**

---

[26] Although this use is common in UK clinical practice, at the time of publication (December 2014]), TCAs, SSRIs and (S)NRIs did not have a UK marketing authorisation for this indication. The prescriber should follow relevant professional guidance, taking full responsibility for the decision. Informed consent should be obtained and documented. See the General Medical Council's Good practice in prescribing and managing medicines and devices for further information.

9.4.2.8 When assessing the risks and benefits of TCAs, SSRIs or (S)NRIs[27] for a woman who is considering breastfeeding, take into account:

- the benefits of breastfeeding for the woman and baby
- the uncertainty about the safety of these drugs for the breastfeeding baby
- the risks associated with switching from or stopping a previously effective medication.

Seek advice from a specialist (preferably from a specialist perinatal mental health service) if there is uncertainty about specific drugs. **[new 2014]**

Benzodiazepines

9.4.2.9 Do not offer benzodiazepines to women in pregnancy and the postnatal period except for the short-term treatment of severe anxiety and agitation. **[2014]**

9.4.2.10 Consider gradually stopping benzodiazepines in women who are planning a pregnancy, pregnant or considering breastfeeding. **[2014]**

Antipsychotic medication

9.4.2.11 When assessing the risks and benefits of antipsychotic medication28 for a pregnant woman, take into account risk factors for gestational diabetes and excessive weight gain. **[new 2014]**

9.4.2.12 When choosing an antipsychotic, take into account that there are limited data on the safety of these drugs in pregnancy and the postnatal period. **[new 2014]**

9.4.2.13 Measure prolactin levels in women who are taking prolactin-raising antipsychotic medication and planning a pregnancy, because raised prolactin levels reduce the chances of conception. If prolactin levels are raised, consider a prolactin-sparing antipsychotic. **[2014]**

9.4.2.14 If a pregnant woman is stable on an antipsychotic and likely to relapse without medication, advise her to continue the antipsychotic. **[new 2014]**

---

[27] Although this use is common in UK clinical practice, at the time of publication (December 2014]), TCAs, SSRIs and (S)NRIs did not have a UK marketing authorisation for this indication. The prescriber should follow relevant professional guidance, taking full responsibility for the decision. Informed consent should be obtained and documented. See the General Medical Council's Good practice in prescribing and managing medicines and devices for further information.

[28] Although this use is common in UK clinical practice, at the time of publication (December 2014]), antipsychotic medication did not have a UK marketing authorisation for this indication. The prescriber should follow relevant professional guidance, taking full responsibility for the decision. Informed consent should be obtained and documented. See the General Medical Council's Good practice in prescribing and managing medicines and devices for further information.

9.4.2.15 Advise pregnant women taking antipsychotic medication about diet and monitor for excessive weight gain, in line with the guideline on weight management before, during and after pregnancy (NICE public health guidance 27). **[new 2014]**

9.4.2.16 Monitor for gestational diabetes in pregnant women taking antipsychotic medication in line with the guideline on diabetes in pregnancy (NICE clinical guideline 63) and offer an oral glucose tolerance test. **[new 2014]**

9.4.2.17 Do not offer depot antipsychotics to a woman who is planning a pregnancy, pregnant or considering breastfeeding, unless she is responding well to a depot and has a previous history of non-adherence with oral medication. **[new 2014]**

Anticonvulsants for mental health problems (valproate, carbamazepine and lamotrigine)

9.4.2.18 Do not offer valproate for acute or long-term treatment of a mental health problem in women who are planning a pregnancy, pregnant or considering breastfeeding. **[new 2014]**

9.4.2.19 If a woman is already taking valproate and is planning a pregnancy, advise her to gradually stop the drug because of the risk of fetal malformations and adverse neurodevelopment outcomes after any exposure in pregnancy. **[2014]**

9.4.2.20 If a woman is already taking valproate and becomes pregnant, stop the drug because of the risk of fetal malformations and adverse neurodevelopmental outcomes. **[2014]**

9.4.2.21 Do not offer carbamazepine to treat a mental health problem in women who are planning a pregnancy, pregnant or considering breastfeeding. **[new 2014]**

9.4.2.22 If a woman is already taking carbamazepine and is planning a pregnancy or becomes pregnant, discuss with the woman the possibility of stopping the drug (because of the risk of adverse drug interactions and fetal malformations). **[new 2014]**

9.4.2.23 If a woman is taking lamotrigine[29] during pregnancy, check lamotrigine levels frequently during pregnancy and into the postnatal period because they vary substantially at these times. **[new 2014]**

---

[29] At the time of publication (December 2014), lamotrigine did not have a UK marketing authorisation for this indication. The prescriber should follow relevant professional guidance, taking full responsibility for the decision. Informed consent should be obtained and documented. See the General Medical Council's Good practice in prescribing and managing medicines and devices for further information.

Lithium

9.4.2.24 Do not offer lithium[30] to women who are planning a pregnancy or pregnant, unless antipsychotic medication has not been effective. **[new 2014]**

9.4.2.25 If antipsychotic medication has not been effective and lithium is offered to a woman who is planning a pregnancy or pregnant, ensure:

- the woman knows that there is a risk of fetal heart malformations when lithium is taken in the first trimester, but the size of the risk is uncertain
- the woman knows that lithium levels may be high in breast milk with a risk of toxicity for the baby
- lithium levels are monitored more frequently throughout pregnancy and the postnatal period. **[new 2014]**

9.4.2.26 If a woman taking lithium becomes pregnant, consider stopping the drug gradually over 4 weeks if she is well. Explain to her that:

- stopping medication may not remove the risk of fetal heart malformations
- there is a risk of relapse, particularly in the postnatal period, if she has bipolar disorder. **[2014]**

9.4.2.27 If a woman taking lithium becomes pregnant and is not well or is at high risk of relapse, consider:

- switching gradually to an antipsychotic or
- stopping lithium and restarting it in the second trimester (if the woman is not planning to breastfeed and her symptoms have responded better to lithium than to other drugs in the past) or
- continuing with lithium if she is at high risk of relapse and an antipsychotic is unlikely to be effective. **[new 2014]**

9.4.2.28 If a woman continues taking lithium during pregnancy:

- check plasma lithium levels every 4 weeks, then weekly from the 36th week
- adjust the dose to keep plasma lithium levels in the woman's therapeutic range
- ensure the woman maintains an adequate fluid balance
- ensure the woman gives birth in hospital

---

[30] Although this use is common in UK clinical practice, at the time of publication (December 2014), lithium did not have a UK marketing authorisation for this indication. The prescriber should follow relevant professional guidance, taking full responsibility for the decision. Informed consent should be obtained and documented. See the General Medical Council's Good practice in prescribing and managing medicines and devices for further information.

- ensure monitoring by the obstetric team when labour starts, including checking plasma lithium levels and fluid balance because of the risk of dehydration and lithium toxicity
- stop lithium during labour and check plasma lithium levels 12 hours after her last dose. **[2014]**

9.5 RECOGNISING MENTAL HEALTH PROBLEMS IN PREGNANCY AND THE POSTNATAL PERIOD AND REFERRAL

9.5.1.1 Recognise that women who have a mental health problem (or are worried that they might have) may be:

- unwilling to disclose or discuss their problem because of fear of stigma, negative perceptions of them as a mother or fear that their baby might be taken into care
- reluctant to engage, or have difficulty in engaging, in treatment because of avoidance associated with their mental health problem or dependence on alcohol or drugs. **[new 2014]**

9.5.1.2 All healthcare professionals referring a woman to a maternity service should ensure that communications with that service (including those relating to initial referral) share information on any past and present mental health problem. **[2014]**

9.5.2 Depression and anxiety disorders

9.5.2.1 Recognise that the range and prevalence of anxiety disorders (including generalised anxiety disorder, obsessive-compulsive disorder, panic disorder, phobias, post-traumatic stress disorder and social anxiety disorder) and depression are under-recognised throughout pregnancy and the postnatal period. **[new 2014]**

9.5.2.2 At a woman's first contact with primary care or her booking visit, and during the early postnatal period, consider asking the following depression identification questions as part of a general discussion about a woman's mental health and wellbeing:

- During the past month, have you often been bothered by feeling down, depressed or hopeless?
- During the past month, have you often been bothered by having little interest or pleasure in doing things?

Also consider asking about anxiety using the 2-item Generalized Anxiety Disorder scale (GAD-2):

- Over the last 2 weeks, how often have you been bothered by feeling nervous, anxious or on edge?[31]
- Over the last 2 weeks, how often have you been bothered by not being able to stop or control worrying?[32] **[new 2014]**

9.5.2.3 If a woman responds positively to either of the depression identification questions in recommendation 9.5.2.2, is at risk of developing a mental health problem, or there is clinical concern, consider:

- using the Edinburgh Postnatal Depression Scale (EPDS) or the Patient Health Questionnaire (PHQ-9) as part of a full assessment or
- referring the woman to her GP or, if a severe mental health problem is suspected, to a mental health professional. **[new 2014]**

9.5.2.4 If a woman scores 3 or more on the GAD-2 scale, consider:

- using the GAD-7 scale for further assessment or
- referring the woman to her GP or, if a severe mental health problem is suspected, to a mental health professional. **[new 2014]**

9.5.2.5 If a woman scores less than 3 on the GAD-2 scale, but you are still concerned she may have an anxiety disorder, ask the following question:

- Do you find yourself avoiding places or activities and does this cause you problems?

If she responds positively, consider:

- using the GAD-7 scale for further assessment or
- referring the woman to her GP or, if a severe mental health problem is suspected, to a mental health professional. **[new 2014]**

9.5.2.6 At all contacts after the first contact with primary care or the booking visit, the health visitor, and other healthcare professionals who have regular contact with a woman in pregnancy and the postnatal period (first year after birth), should consider:

- asking the 2 depression identification questions and the GAD-2 (see recommendation 9.5.2.2) as part of a general discussion about her mental health and wellbeing and
- using the EPDS or the PHQ-9 as part of monitoring. **[new 2014]**

9.5.3 Severe mental illness

---

[31] An answer of 'Not at all' scores 0; 'Several days' scores 1; 'More than half the days' scores 2; 'Nearly every day' scores 3.

[32] An answer of 'Not at all' scores 0; 'Several days' scores 1; 'More than half the days' scores 2; 'Nearly every day' scores 3.

9.5.3.1 At a woman's first contact with services in pregnancy and the postnatal period, ask about:

- any past or present severe mental illness
- past or present treatment by a specialist mental health service, including inpatient care
- any severe perinatal mental illness in a first-degree relative (mother, sister or daughter). **[2014]**

9.5.3.2 Refer to a secondary mental health service (preferably a specialist perinatal mental health service) for assessment and treatment, all women who:

- have or are suspected to have severe mental illness
- have any history of severe mental illness (during pregnancy or the postnatal period or at any other time).

Ensure that the woman's GP knows about the referral. **[new 2014]**

9.5.3.3 If a woman has any past or present severe mental illness or there is a family history of severe perinatal mental illness in a first-degree relative, be alert for possible symptoms of postpartum psychosis in the first 2 weeks after childbirth. **[new 2014]**

9.5.3.4 If a woman has sudden onset of symptoms suggesting postpartum psychosis, refer her to a secondary mental health service (preferably a specialist perinatal mental health service) for immediate assessment (within 4 hours of referral). **[new 2014]**

9.5.4 Alcohol and drug misuse

9.5.4.1 If alcohol misuse is suspected, use the Alcohol Use Disorders Identification Test (AUDIT) as an identification tool in line with recommendation 1.2.1.4 of the guideline on alcohol-use disorders (NICE clinical guideline 115). **[new 2014]**

9.5.4.2 If drug misuse is suspected, follow the recommendations on identification and assessment in section 1.2 of the guideline on drug misuse – psychosocial interventions (NICE clinical guideline 51). **[new 2014]**

9.6 ASSESSMENT AND CARE PLANNING IN PREGNANCY AND THE POSTNATAL PERIOD

9.6.1.1 Assessment and diagnosis of a suspected mental health problem in pregnancy and the postnatal period should include:

- history of any mental health problem, including in pregnancy or the postnatal period

- physical wellbeing (including weight, smoking, nutrition and activity level) and history of any physical health problem
- alcohol and drug misuse
- the woman's attitude towards the pregnancy, including denial of pregnancy
- the woman's experience of pregnancy and any problems experienced by her, the fetus or the baby
- the mother–baby relationship
- any past or present treatment for a mental health problem, and response to any treatment
- social networks and quality of interpersonal relationships
- living conditions and social isolation
- family history (first-degree relative) of mental health problems
- domestic violence and abuse, sexual abuse, trauma or childhood maltreatment
- housing, employment, economic and immigration status
- responsibilities as a carer for other children and young people or other adults. **[new 2014]**

9.6.1.2 When assessing or treating a mental health problem in pregnancy or the postnatal period, take account of any learning disabilities or acquired cognitive impairments, and assess the need to consult with a specialist when developing care plans. **[new 2014]**

9.6.1.3 Carry out a risk assessment in conjunction with the woman and, if she agrees, her partner, family or carer. Focus on areas that are likely to present possible risk such as self-neglect, self-harm, suicidal thoughts and intent, risks to others (including the baby), smoking, drug or alcohol misuse and domestic violence and abuse. **[new 2014]**

9.6.1.4 If there is a risk of, or there are concerns about, suspected child maltreatment, follow local safeguarding protocols. **[new 2014]**

9.6.1.5 If there is a risk of self-harm or suicide:

- assess whether the woman has adequate social support and is aware of sources of help
- arrange help appropriate to the level of risk
- inform all relevant healthcare professionals (including the GP and those identified in the care plan [see recommendation 9.6.1.6])
- advise the woman, and her partner, family or carer, to seek further help if the situation deteriorates. **[new 2014]**

9.6.1.6 Professionals in secondary mental health services, including specialist perinatal mental health services, should develop a written care plan in collaboration with a woman who has or has had a severe mental illness. If she agrees, her partner, family or carer should also be involved. The plan

should cover pregnancy, childbirth and the postnatal period (including the potential impact of the illness on the baby) and should include:

- a clear statement of jointly agreed treatment goals and how outcomes will be routinely monitored
- increased contact with and referral to specialist perinatal mental health services
- the names and contact details of key professionals.

The care plan should be recorded in all versions of the woman's notes (her own records and maternity, primary care and mental health notes) and a copy given to the woman and all involved professionals. **[new 2014]**

9.7 PROVIDING INTERVENTIONS IN PREGNANCY AND THE POSTNATAL PERIOD

9.7.1.1 All healthcare professionals providing assessment and interventions for mental health problems in pregnancy and the postnatal period should understand the variations in their presentation and course at these times, how these variations affect treatment, and the context in which they are assessed and treated (for example, maternity services, health visiting and mental health services). **[new 2014]**

9.7.1.2 All interventions for mental health problems in pregnancy and the postnatal period should be delivered by competent practitioners. Psychological and psychosocial interventions should be based on the relevant treatment manual(s), which should guide the structure and duration of the intervention. Practitioners should consider using competence frameworks developed from the relevant treatment manual(s) and for all interventions practitioners should:

- receive regular high-quality supervision
- use routine outcome measures and ensure that the woman is involved in reviewing the efficacy of the treatment
- engage in monitoring and evaluation of treatment adherence and practitioner competence – for example, by using video and audio tapes, and external audit and scrutiny where appropriate.

[new 2014][33]

9.7.1.3 When a woman with a known or suspected mental health problem is referred in pregnancy or the postnatal period, assess for treatment within 2 weeks of referral and provide psychological interventions within 1 month of initial assessment. **[new 2014]**

9.7.1.4 When offering psychotropic medication during pregnancy and the postnatal period, follow the principles in recommendations 9.4.2.1 to 9.4.2.28. **[new 2014]**

9.7.1.5 Provide interventions for mental health problems in pregnancy and the postnatal period within a stepped-care model of service delivery in line with recommendation 1.5.1.3 of the guideline on common mental health disorders (NICE clinical guideline 123). **[new 2014]**

[33] Adapted from the guideline on depression in adults (NICE clinical guideline 90).

9.8 TREATING SPECIFIC MENTAL HEALTH PROBLEMS IN PREGNANCY AND THE POSTNATAL PERIOD

9.8.1 Interventions for depression

9.8.1.1 For a woman with persistent subthreshold depressive symptoms, or mild to moderate depression, in pregnancy or the postnatal period, consider facilitated self-help (delivered as described in recommendation 1.4.2.2 of the guideline on depression in adults [NICE clinical guideline 90]). **[new 2014]**

9.8.1.2 For a woman with a history of severe depression who initially presents with mild depression in pregnancy or the postnatal period, consider a TCA, SSRI or (S)NRI. **[new 2014]**

9.8.1.3 For a woman with moderate or severe depression in pregnancy or the postnatal period, consider the following options:

- a high-intensity psychological intervention (for example, CBT)
- a TCA, SSRI or (S)NRI if the woman understands the risks associated with the medication and the mental health problem in pregnancy and the postnatal period and:
  - she has expressed a preference for medication or
  - she declines psychological interventions or
  - her symptoms have not responded to psychological interventions
- a high-intensity psychological intervention in combination with medication if the woman understands the risks associated with the medication and the mental health problem in pregnancy and the postnatal period and there is no response, or a limited response, to a high-intensity psychological intervention or medication alone. **[new 2014]**

9.8.1.4 If a woman who is taking a TCA, SSRI or (S)NRI for mild to moderate depression becomes pregnant, discuss stopping the medication gradually and consider facilitated self-help (delivered as described in recommendation 1.4.2.2 of the guideline on depression in adults [NICE clinical guideline 90]). **[new 2014]**

9.8.1.5 If a pregnant woman is taking a TCA, SSRI or (S)NRI for moderate depression and wants to stop her medication, take into account previous response to treatment, stage of pregnancy, risk of relapse, risk associated with medication and her preference, and discuss with her the following options:

- switching to a high-intensity psychological intervention (for example, CBT)
- changing medication if there is a drug that is effective for her with a lower risk of adverse effects. **[new 2014]**

9.8.1.6 If a pregnant woman is taking a TCA, SSRI or (S)NRI for severe depression, take into account previous response to treatment, stage of pregnancy, risk of relapse, risk associated with medication and her preference, and discuss with her the following options:

- continuing with the current medication
- changing medication if there is a drug that is effective for her with a lower risk of adverse effects
- combining medication with a high-intensity psychological intervention (for example, CBT)
- switching to a high-intensity psychological intervention (for example, CBT) if she decides to stop taking medication. **[new 2014]**

9.8.2 Interventions for anxiety disorders

9.8.2.1 For a woman with tokophobia (an extreme fear of childbirth), offer an opportunity to discuss her fears with a healthcare professional with expertise in providing perinatal mental health support in line with section 1.2.9 of the guideline on caesarean section (NICE clinical guideline 132). **[new 2014]**

9.8.2.2 For a woman with persistent subthreshold symptoms of anxiety in pregnancy or the postnatal period, consider facilitated self-help. This should consist of use of CBT-based self-help materials over 2–3 months with support (either face to face or by telephone) for a total of 2–3 hours over 6 sessions. **[new 2014]**

9.8.2.3 For a woman with an anxiety disorder in pregnancy or the postnatal period, offer a low-intensity psychological intervention (for example, facilitated self-help) or a high-intensity psychological intervention (for example, CBT) as initial treatment in line with the recommendations set out in the NICE guideline for the specific mental health problem and be aware that:

- only high-intensity psychological interventions are recommended for post-traumatic stress disorder
- high-intensity psychological interventions are recommended for the initial treatment of social anxiety disorder
- progress should be closely monitored and a high-intensity psychological intervention offered within 2 weeks if symptoms have not improved. **[new 2014]**

9.8.2.4 If a woman who is taking a TCA, SSRI or (S)NRI for an anxiety disorder becomes pregnant, discuss with her the following options:

- stopping the medication gradually and switching to a high-intensity psychological intervention (for example, CBT)

- continuing with medication if she understands the risks associated with the medication and the mental health problem in pregnancy and the postnatal period and:
  - has expressed a preference for medication or
  - declines psychological interventions or
  - her symptoms have not responded to psychological interventions
- changing medication if there is a drug that is effective for her with a lower risk of adverse effects
- combining medication with a high-intensity psychological intervention (for example, CBT) if the woman understands the risks associated with the medication and the mental health problem in pregnancy and the postnatal period and there is no response, or a limited response, to a high-intensity psychological intervention alone. **[new 2014]**

9.8.3 Psychological interventions for eating disorders

9.8.3.1 For a woman with an eating disorder in pregnancy or the postnatal period:

- offer a psychological intervention in line with the guideline on eating disorders (NICE clinical guideline 9)
- monitor the woman's condition carefully throughout pregnancy and the postnatal period
- assess the need for fetal growth scans
- discuss the importance of healthy eating during pregnancy and the postnatal period in line with the guideline on maternal and child nutrition (NICE public health guidance 11)
- advise her about feeding the baby in line with the guideline on maternal and child nutrition (NICE public health guidance 11) and support her with this. **[new 2014]**

9.8.4 Interventions for alcohol and drug misuse

9.8.4.1 If hazardous drug or alcohol misuse is identified in pregnancy or the postnatal period, refer or offer brief interventions in line with section 1.3.1 of the guideline on drug misuse – psychosocial interventions (NICE clinical guideline 51) or the guideline on alcohol-use disorders: preventing harmful drinking (NICE public health guidance 24). **[new 2014]**

9.8.4.2 If harmful or dependent drug or alcohol misuse is identified in pregnancy or the postnatal period, refer the woman to a specialist substance misuse service for advice and treatment. **[new 2014]**

9.8.4.3 Offer assisted alcohol withdrawal in collaboration with specialist mental health and alcohol services (preferably in an inpatient setting) to pregnant women who are dependent on alcohol. Work with a woman who does not want assisted alcohol withdrawal to help her reduce her alcohol intake. **[new 2014]**

9.8.4.4 Offer detoxification in collaboration with specialist mental health and substance misuse services to pregnant women who are dependent on opioids. Monitor closely after completion of detoxification. Work with a woman who does not want detoxification to help her reduce her opioid intake. Recognise the risk of accidental overdose in women who stop or reduce drug misuse in pregnancy but start misusing again after childbirth. **[new 2014]**

9.8.5 Interventions for severe mental illness

9.8.5.1 Consider psychological interventions for women with bipolar disorder. This includes:

- CBT, IPT and behavioural couples therapy for bipolar depression
- structured individual, group and family interventions designed for bipolar disorder to reduce the risk of relapse, particularly when medication is changed or stopped. **[new 2014]**

9.8.5.2 If a pregnant woman develops mania or psychosis and is not taking psychotropic medication, offer an antipsychotic. **[new 2014]**

9.8.5.3 Consider psychological interventions (CBT or family intervention) delivered as described in section 1.3.7 of the guideline on psychosis and schizophrenia in adults (NICE clinical guideline 178) for a woman with psychosis or schizophrenia who becomes pregnant and is at risk of relapse arising from:

- stress associated with pregnancy or the postnatal period
- a change in medication, including stopping antipsychotic medication. **[new 2014]**

9.8.5.4 Offer an antipsychotic in line with recommendations 1.5.3 and 1.5.4 of the guideline on bipolar disorder (NICE clinical guideline 185) as prophylactic medication if a woman with bipolar disorder:

- becomes pregnant and is stopping lithium, or
- plans to breastfeed. **[new 2014]**

9.8.5.5 If a pregnant woman with bipolar disorder develops mania while taking prophylactic medication:

- check the dose of the prophylactic medication and adherence
- increase the dose if the prophylactic medication is an antipsychotic

- suggest changing to an antipsychotic if she is taking another type of prophylactic medication
- consider lithium if there is no response to an increase in dose or change of drug and the woman has severe mania
- consider electroconvulsive therapy (ECT) if there has been no response to lithium. **[new 2014]**

9.8.6 Interventions for sleep problems

9.8.6.1 Advise pregnant women who have a sleep problem about sleep hygiene (including having a healthy bedtime routine, avoiding caffeine and reducing activity before sleep). For women with a severe or chronic sleep problem, consider promethazine[34]. **[new 2014]**

9.8.7 Electroconvulsive therapy

9.8.7.1 Consider electroconvulsive therapy (ECT) for pregnant women with severe depression, severe mixed affective states or mania, or catatonia, whose physical health or that of the fetus is at serious risk. **[2014]**

[34] At the time of publication (December 2014), promethazine did not have a UK marketing authorisation for this indication. The prescriber should follow relevant professional guidance, taking full responsibility for the decision. Informed consent should be obtained and documented. See the General Medical Council's Good practice in prescribing and managing medicines and devices for further information.

9.8.8 Rapid tranquillisation

9.8.8.1 A pregnant woman requiring rapid tranquillisation should be treated according to the NICE clinical guidelines on the short-term management of disturbed/violent behaviour, schizophrenia and bipolar disorder (see the related NICE guidance in section 3.2 for details [in the NICE guideline]), except that:

- she should not be secluded after rapid tranquillisation
- restraint procedures should be adapted to avoid possible harm to the fetus
- when choosing an agent for rapid tranquillisation in a pregnant woman, an antipsychotic or a benzodiazepine with a short half-life should be considered; if an antipsychotic is used, it should be at the minimum effective dose because of neonatal extrapyramidal symptoms; if a benzodiazepine is used, the risks of floppy baby syndrome should be taken into account
- during the perinatal period, the woman's care should be managed in close collaboration with a paediatrician and an anaesthetist. **[2007]**

9.9 CONSIDERATIONS FOR WOMEN AND THEIR BABIES IN THE POSTNATAL PERIOD

9.9.1 Reviewing treatment for women with severe mental illness

9.9.1.1 After childbirth, review and assess the need for starting, restarting or adjusting psychotropic medication as soon as a woman with a past or present severe mental illness is medically stable. **[new 2014]**

9.9.2 Monitoring babies for effects of psychotropic medication taken in pregnancy

9.9.2.1 If a woman has taken psychotropic medication during pregnancy, carry out a full neonatal assessment of the newborn baby, bearing in mind:

- the variation in the onset of adverse effects of psychotropic medication
- the need for further monitoring
- the need to inform relevant healthcare professionals and the woman and her partner, family or carer of any further monitoring, particularly if the woman has been discharged early. **[new 2014]**

9.9.3 Care of women and their babies if there has been alcohol or drug misuse in pregnancy

9.9.3.1 If there has been alcohol or drug misuse in pregnancy, offer treatment and support after childbirth to both the woman and the baby, including:

- a full neonatal assessment for any congenital abnormalities or neonatal adaptation syndrome
- continuing psychological treatment and support for the woman
- monitoring of the baby. **[new 2014]**

9.9.4 Traumatic birth, stillbirth and miscarriage

9.9.4.1 Offer advice and support to women who have had a traumatic birth or miscarriage and wish to talk about their experience. Take into account the effect of the birth or miscarriage on the partner and encourage them to accept support from family and friends. **[new 2014]**

9.9.4.2 Offer women who have post-traumatic stress disorder, which has resulted from a traumatic birth, miscarriage, stillbirth or neonatal death, a high-intensity psychological intervention (trauma-focused CBT or eye movement desensitisation and reprocessing [EMDR]) in line with the guideline on post-traumatic stress disorder (PTSD) (NICE clinical guideline 26). **[new 2014]**

9.9.4.3 Do not offer single-session high-intensity psychological interventions with an explicit focus on 're-living' the trauma to women who have a traumatic birth. **[new 2014]**

9.9.4.4 Discuss with a woman whose baby is stillborn or dies soon after birth, and her partner and family, the option of 1 or more of the following:

- seeing a photograph of the baby
- having mementos of the baby
- seeing the baby
- holding the baby.

This should be facilitated by an experienced practitioner and the woman and her partner and family should be offered a follow-up appointment in primary or secondary care. If it is known that the baby has died in utero, this discussion should take place before the delivery, and continue after delivery if needed. **[new 2014]**

9.9.5 Psychotropic medication and breastfeeding

9.9.5.1 Encourage women with a mental health problem to breastfeed, unless they are taking carbamazepine, clozapine or lithium (valproate is not recommended to treat a mental health problem in women of childbearing potential). However, support each woman in the choice of feeding method that best suits her and her family. **[new 2014]**

9.9.5.2 When assessing the risks and benefits of TCAs, SSRIs or (S)NRIs for women who are breastfeeding, take into account:

- the limited data about the safety of these drugs and
- the risks associated with switching from a previously effective medication.

Seek advice from a specialist (preferably from a specialist perinatal mental health service) if needed for specific drugs. **[new 2014]**

9.9.5.3 When assessing the risks and benefits of antipsychotic medication for women who are breastfeeding, take into account:

- the limited data on the safety of these drugs and
- the level of antipsychotic medication in breast milk depends on the drug. **[new 2014]**

9.9.5.4 If a woman is taking psychotropic medication while breastfeeding, monitor the baby for adverse effects. **[2014]**

9.9.6 The mother–baby relationship

9.9.6.1 Recognise that some women with a mental health problem may experience difficulties with the mother–baby relationship. Assess the nature of this relationship, including verbal interaction, emotional sensitivity and physical care, at all postnatal contacts. Discuss any concerns that the woman has about her relationship with her baby and provide information and treatment for the mental health problem. **[new 2014]**

9.9.6.2 Consider further intervention to improve the mother–baby relationship if any problems in the relationship have not resolved. **[new 2014]**

9.10 THE ORGANISATION OF SERVICES

9.10.1.1 Women who need inpatient care for a mental health problem within 12 months of childbirth should normally be admitted to a specialist mother and baby unit, unless there are specific reasons for not doing so. **[2007]**

9.10.1.2 Managers and senior healthcare professionals responsible for perinatal mental health services (including those working in maternity and primary care services) should ensure that:

- there are clearly specified care pathways so that all primary and secondary healthcare professionals involved in the care of women during pregnancy and the postnatal period know how to access assessment and treatment
- staff have supervision and training, covering mental health problems, assessment methods and referral routes, to allow them to follow the care pathways. **[2007]**

9.10.1.3 Clinical networks should be established for perinatal mental health services, managed by a coordinating board of healthcare professionals, commissioners, managers, and service users and carers. These networks should provide:

- a specialist multidisciplinary perinatal service in each locality, which provides direct services, consultation and advice to maternity services, other mental health services and community services; in areas of high morbidity these services may be provided by separate specialist perinatal teams
- access to specialist expert advice on the risks and benefits of psychotropic medication during pregnancy and breastfeeding
- clear referral and management protocols for services across all levels of the existing stepped-care frameworks for mental health problems, to ensure effective transfer of information and continuity of care
- pathways of care for service users, with defined roles and competencies for all professional groups involved. **[2007]**

9.10.1.4 Each managed perinatal mental health network should have designated specialist inpatient services and cover a population where there are between 25,000 and 50,000 live births a year, depending on the local psychiatric morbidity rates. **[2007]**

9.10.1.5 Specialist perinatal inpatient services should:

- provide facilities designed specifically for mothers and babies (typically with 6–12 beds)
- be staffed by specialist perinatal mental health staff
- be staffed to provide appropriate care for babies
- have effective liaison with general medical and mental health services
- have available the full range of therapeutic services
- be closely integrated with community-based mental health services to ensure continuity of care and minimum length of stay. **[2007]**

10 APPENDICES

All appendices are in separate files on the NICE and NCCMH websites.

No table of contents entries found.Appendix 17: Clinical evidence – completed methodology checklists

Appendix 18: Clinical evidence – study characteristics tables

Appendix 19: Clinical evidence – forest plots

Appendix 20: Economic evidence – completed methodology checklists

Appendix 21: Economic evidence – evidence tables of published studies

Appendix 22: Clinical and economic evidence profiles

Appendix 23: 2007 Methods chapter

Appendix 24: 2007 Health economic evidence on mother and baby units

Appendix 25: 2007 Survey of antenatal and postnatal mental health primary care services in England and Wales – questionnaire

Appendix 26: 2007 Survey results of antenatal and postnatal mental health primary care services in England and Wales – questionnaire

Appendix 27: 2007 Chapter 4 (Population, Disorders and Services)

11 REFERENCES

Ackerman JP, Llorente AM, Black MM, Ackerman CS, Mayes LA, Nair P. The effect of prenatal drug exposure and caregiving context on children's performance on a task of sustained visual attention. Journal of Developmental and Behavioral Pediatrics. 2008;29:467-74.

Adab N, Jacoby A, Smith D, Chadwick D. Additional educational needs in children born to mothers with epilepsy. Journal of Neurology, Neurosurgery and Psychiatry. 2001;70:15-21.

Adab N, Kini U, Vinten J, Ayres J, Baker G, Clayton-Smith J, et al. The longer term outcome of children born to mothers with epilepsy. Journal of Neurology, Neurosurgery and Psychiatry. 2004;75:1575-83.

Adewuya AO, Eegunranti AB, Lawal AM. Prevalence of postnatal depression in Western Nigerian women: a controlled study. International Journal of Psychiatry in Clinical Practice. 2005;9:60-4.

Adewuya AO, Ola BA, Dada AO, Fasoto OO. Validation of the Edinburgh Postnatal Depression Scale as a screening tool for depression in late pregnancy among Nigerian women. Journal of Psychosomatic Obstetrics and Gynecology. 2006;27:267-72.

Adeyemi A, Mosaku K, Ajenifuja O, et al. Depressive symptoms in a sample of women following perinatal loss. Journal of the National Medical Association. 2008;100:1463-8.

Agoub M, Moussaoui D, Battas O. Prevalence of postpartum depression in a Moroccan sample. Archives of Women's Mental Health. 2005;8:37-43.

AGREE Collaboration. Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: the AGREE project. Quality and Safety in Health Care. 2003;12:18-23.

Almgren M, Kallen B, Lavebratt C. Population-based study of antiepileptic drug exposure in utero-Influence on head circumference in newborns. Seizure. 2009;18:672-75.

Altman D, Bland M. Statistics notes: diagnostic tests 2: predictive values. British Medical Journal. 1994a;309:102.

Altman D, Bland M. Statistics notes: diagnostic tests 1: sensitivity and specificity. British Medical Journal. 1994b;308:1552. Alvarado-Esquivel C, Sifuentes-Alvarez A, Salas-Martinez C, Martínez-García S. Validation of the Edinburgh Postpartum Depression Scale in a population of puerperal women in Mexico. Clinical Practice and Epidemiology in Mental Health. 2006;2:33.

Alwan S, Reefhuis J, Rasmussen SA, Olney RS, Friedman JM. Use of selective serotonin-reuptake inhibitors in pregnancy and the risk of birth defects. New England Journal of Medicine. 2007;356:2684-92.

Amato L, Davoli M, Minozzi S, Ferroni E, Ali R, Ferri M. Methadone at tapered doses for the management of opioid withdrawal (Review). Cochrane Database of Systematic Reviews. 2013;2:Art. No.: CD003409. DOI: 10.1002/14651858.CD003409.pub4

Amendah DD, Grosse SD, Bertrand J. Medical expenditures of children in the United States with fetal alcohol syndrome. Neurotoxicology and Teratology. 2011;33:322-4.

American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (5th edn). Arlington, VA: American Psychiatric Publishing; 2013.

Ammerman RT, Putnam FW, Altaye M, Stevens J, Teeters AR, Van Ginkel JB. A clinical trial of in-home CBT for depressed mothers in home visitation. Behaviour Therapy. 2013a;44:359-72.

Ammerman RT, Putnam FW, Altaye M, Teeters AR, Stevens J, Van Ginkel JB. Treatment of depressed mothers in home visiting: impact on psychological distress and social functioning. Child Abuse and Neglect. 2013b;37:544-54.

Andersen LB, Melvaer LB, Videbech P, Lamont RF, Joergensen JS. Risk factors for developing post-traumatic stress disorder following childbirth: a systematic review. Acta Obstetricia et Gynecologica Scandinavica. 2012;91:1261-72.

Antonysamy A, Wieck A, Wittkowski A. Service satisfaction on discharge from a psychiatric mother and baby unit: a representative patient survey. Archives of Women's Mental Health. 2009;12:359-62.

Appleby L, Warner R, Whitton A, Faragher B. A controlled study of fluoxetine and cognitive-behavioural counselling in the treatment of postnatal depression. British Medical Journal. 1997;314:932-36.

Aracena M, Krause M, Perez C, Mendez MJ, Salvatierra L, Soto M, et al. A cost-effectiveness evaluation of a home visit program for adolescent mothers. Journal of Health Psychology. 2009;14:878-887.

Areias ME, Kumar R, Barros H, Figueiredo E. Comparative incidence of depression in women and men, during pregnancy and after childbirth. Validation of the Edinburgh Postnatal Depression Scale in Portuguese mothers. British Journal of Psychiatry. 1996;169:30-5.

Armstrong KL, Fraser JA, Dadds MR, Morris J. A randomized, controlled trial of nurse home visiting to vulnerable families with newborns. Journal of Paediatric Child Health. 1999;35:237-44.

Armstrong KL, Fraser JA, Dadds MR, Morris J. Promoting secure attachment, maternal mood and child health in a vulnerable population: a randomized controlled trial. Journal of Paediatric Child Health. 2000;36:555-62.

Armstrong K, Edwards H. The effects of exercise and social support on mothers reporting depressive symptoms: a pilot randomized controlled trial. International Journal of Mental Health Nursing. 2003;12:130-38.

Armstrong K, Edwards H. The effectiveness of a pram-walking exercise programme in reducing depressive symptomatology for postnatal women. International Journal of Nursing Practice. 2004;10:177-94.

Armstrong MA, Kaskutas LA, Witbrodt J, Taillac CJ, Hung YY, Osejo VM, et al. Using drink size to talk about drinking during pregnancy: a randomized clinical trial of early start plus. Social Work in Health Care. 2009;48:90-103.

Arroll B, Goodyear-Smith F, Kerse N, Fishman T, Gunn J. 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. 2005;331:884

Artama M, Auvinen A, Raudaskoski T, Isojarvi I, Isojarvi J. Antiepileptic drug use of women with epilepsy and congenital malformations in offspring. Neurology. 2005:64;1874-8.

Artama M, Gissler M, Malm H, Ritvanen A. Effects of maternal epilepsy and antiepileptic drug use during pregnancy on perinatal health in offspring: Nationwide, retrospective cohort study in Finland. Drug Safety. 2013:36;359-69

AscasoTerren C, Garcia Esteve L, Navarro P, Aguado J, Ojuel J, Tarragona MJ. Prevalence of postpartum depression in Spanish mothers: comparison of estimation by mean of the structured clinical interview for DSM-IV with the Edinburgh Postnatal Depression Scale. MedicinaClinica. 2003;120:326-9.

Auerbach JG, Hans SL, Marcus J, Maeir S. Maternal psychotropic medication and neonatal behavior. Neurotoxicology & Teratology. 1992;14:399-406.

Austin M-P, Frilingos M, Lumley J, Hadzi-Pavlovic D, Roncolato W, Acland S, et al. Brief antenatal cognitive behaviour therapy group intervention for the prevention of postnatal depression and anxiety: a randomised controlled trial. Journal of Affective Disorders. 2008;105:35-44.

Aydin N, Inandi T, Yigit A, Hodoglugil NN.Validation of the Turkish version of the Edinburgh Postnatal Depression Scale among women within their first postpartum year. Social Psychiatry and Psychiatric Epidemiology. 2004;39:483-6.

Ayers S, Eagle A, Waring H. The effects of childbirth-related post-traumatic stress disorder on women and their relationships: a qualitative study. Psychology, Health and Medicine. 2006;11:389-98.

Backes CH, Backes CR, Gardner D, Nankervis CA, Giannone PJ, Cordero L. Neonatal abstinence syndrome (NAS): transitioning methadone treated infants from an inpatient to an outpatient setting. Journal of Perinatology. 2012;32:425-30.

Baggaley RF, Ganaba R, Filippi V, Kere M, Marshall T, Sombie I, Storeng KT, Patel V. Detecting depression after pregnancy: the validity of the K10 and K6 in Burkina Faso. Tropical Medicine and International Health. 2007;12:1225-1229.

Bakker MK, Kerstjens-Frederikse WS, Buys CHCM, De Walle HEK, De Jong-van Den Berg LTW. First-trimester use of paroxetine and congenital heart defects: A population-based case-control study. Birth Defects Research Part A - Clinical and Molecular Teratology. 2010;88:94-100.

Ban L, Gibson JE, West J, Fiaschi L, Oates MR, Tata LJ. Impact of socioeconomic deprivation on maternal perinatal mental illnesses presenting to UK general practice. British Journal of General Practice. 2012;62:671-8.

Ban L, West J, Gibson JE, Fiaschi L, Sokal R, Doyle P, et al. First trimester exposure to anxiolytic and hypnotic drugs and the risks of major congenital anomalies: a United Kingdom popluation-based cohort study. PLoS ONE. 2014;9:e100996.

Barker ED, Jaffee SR, Uher R, Maughan B. The contribution of prenatal and postnatal maternal anxiety and depression to child maladjustment. Depression and Anxiety. 2011;28:696-702.

Barlow J, Davis H, McIntosh E, Jarrett P, Mockford C, Stewart-Brown S. Role of home visiting in improving parenting and health in families at risk of abuse and neglect: results of a multicentre randomised controlled trial and economic evaluation. Archives of Disease Childhood. 2007;92:229-33.

Barnet B, Liu J, DeVoe M, Alperovitz-Bichell K, Duggan AK. Home visiting for adolescent mothers: effects on parenting, maternal life course, and primary care linkage. Annals of Family Medicine. 2007;5:224-32.

Barnett B, Matthey S, Gyaneshwar R. Screening for postnatal depression in women of non-English speaking background. Archives of Women's Mental Health. 1999;2:67-74.

Bartu A, Sharp J, Ludlow J, Doherty DA. Postnatal home visiting for illicit drug-using mothers and their infants: a randomised controlled trial. Australian and New Zealand Journal of Obstetrics and Gynaecology. 2006;46:419-26.

Beardslee WR, Bemporad J, Keller MB, Klerman GL. Children of parents with major affective disorder: a review. The American Journal of Psychiatry. 1983;140:825-32.

Beck CT, Gable RK. Comparative analysis of the performance of the Postpartum Depression Screening Scale with two other depression instruments. Nursing Research. 2001;50:242-250.

Bennedsen BE, Mortensen PB, Olesen A, Henriksen TB. Congenital malformations, stillbirths, and infant deaths among children of women with schizophrenia. Archives of General Psychiatry. 2001;58:674-9.

Benvenuti P, Ferrara M, Niccolai C, Valoriani V, Cox JL. The Edinburgh Postnatal Depression Scale: Validation for an Italian sample. Journal of Affective Disorders. 1999;53:137-41.

Bergink V, Kooistra L, Lambregtse-van den Berg MP, Wijnen H, Bunevicius R, van Baar A, Pop V. Validation of the Edinburgh Depression Scale during pregnancy.Journal of Psychosomatic Research. 2011;70:385-9.

Berle JO, Aarre TF, Mykletun A, Dahl AA, Holsten F. Screening for postnatal depression: Validation of the Norwegian version of the Edinburgh Postnatal Depression Scale, and assessment of risk factors for postnatal depression. Journal of Affective Disorders. 2003;76:151-56.

Berlin JA. Does blinding of readers affect the results of meta-analyses? Lancet. 2001;350:185-6.

Bernard RS, Williams SE, Storfer-Isser A, Rhine W, Horwitz SM, Koopman C, et al. Brief cognitive-behavioral intervention for maternal depression and trauma in the neonatal intensive care unit: a pilot study. Journal of Traumatic Stress. 2011;24:230-4.

Bilszta JLC, Buist AE, Wang F, Zulkefli NR. Use of video feedback intervention in an inpatient perinatal psychiatric setting to improve maternal parenting. Archives of Women's Mental Health. 2012;15:249-57.

Black MM, Nair P, Kight C, Wachtel R, Roby P, Schuler M. Parenting and early development among children of drug-abusing women: effects of home intervention. Pediatrics. 1994;94:440-8.

Blair MM, Glynn LM, Sandman CA, Davis EP. Prenatal maternal anxiety and early childhood temperament. Stress. 2011;14:644-51.

Bloch M, Schmidt PJ, Danaceau M, Murphy J, Nieman L, Rubinow DR. Effects of gonadal steroids in women with a history of postpartum depression. American Journal of Psychiatry. 2000;157:924-30.

Bloch M, Meiboom H, Lorberblatt M, Bluvstein I, Aharonov I, Schreiber S. The effect of sertraline add-on to brief dynamic psychotherapy for the treatment of postpartum depression: A randomized, double-blind, placebo-controlled study. Journal of Clinical Psychiatry. 2012;73:235-41.

Boath E, Major K, Cox J. When the cradle falls II: The cost-effectiveness of treating postnatal depression in a psychiatric day hospital compared with routine primary care. Journal of Affective Disorders. 2003;74:159-66.Boath E, Bradley E, Henshaw C. Women's views of antidepressants in the treatment of postnatal depression. Journal of Psychosomatic Obstetrics and Gynecology. 2004;25:221-33.

Boden R, Lundgren M, Brandt L, Reutfors J, Andersen M, Kieler H. Risks of adverse pregnancy and birth outcomes in women treated or not treated with mood stabilisers for bipolar disorder: Population based cohort study. BMJ (Online). 2012a;345.

Boden R, Lundgren M, Brandt L, Reutfors J, Kieler H. Antipsychotics during pregnancy: relation to fetal and maternal metabolic effects. Archives of General Psychiatry. 2012b;96:715-21.

Bonnot O, Vollset SE, Godet PF, D'Amato T, Robert E. Maternal exposure to lorazepam and anal atresia in newborns: Results from a hypothesis-generating study of benzodiazepines and malformations. Journal of Clinical Psychopharmacology. 2001;21:456-58.

Booth A. Chapter 6: Formulating research questions. In: Booth A, Brice A, eds. Evidence based practice: A handbook for information professionals. London: Facet; 2003.

Borjesson K, Ruppert S, Bagedahl-Strindlund M. A longitudinal study of psychiatric symptoms in primiparous women: relation to personality disorders and sociodemographic factors. Archives of Women's Mental Health. 2005;8:232-42.

Borthen I, Eide MG, Daltveit AK, Gilhus NE. Obstetric outcome in women with epilepsy: a hospital-based, retrospective study. BJOG: an International Journal of Obstetrics and Gynaecology. 2011;118:956-65.

Boucher N, Bairam A, Beaulac-Baillargeon L. A new look at the neonate's clinical presentation after in utero exposure to antidepressants in late pregnancy. Journal of clinical psychopharmacology. 2008:28;334-9.

Boyce P, Stubbs J, Todd A. The Edinburgh Postnatal Depression Scale: validation for an Australian sample. Australian and New Zealand Journal of Psychiatry. 1993;27:472-476.

Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. Journal of Health Economics. 2002;21:271-292

Breustedt S, Puckering C. A qualitative evaluation of women's experiences of the mellow bumps antenatal intervention. British Journal of Midwifery. 2013;21:187-194.

Briggs A, Sculpher M, Claxton K. Decision Modelling for Health Economic Evaluation. New York: 2006.

Brinch M, Isager T, Tolstrup K. Anorexia nervosa and motherhood: reproduction pattern and mothering behavior of 50 women. Acta Psychiatrica Scandinavica. 1988;77:611-17.

Brockington IF. Motherhood and Mental Health. Oxford: Oxford University Press; 1996.

Bromley RL, Mawer GE, Briggs M, Cheyne C, Clayton-Smith J, García-Fiñana et al. The prevalence of neurodevelopmental disorders in children prenatally exposed to antiepileptic drugs. Journal of Neurology, Neurosurgery and Psychiatry. 2013;84:637-643.

Brosh K, Matok I, Sheine E, Koren G, Wiznitzer A, Gorodischer R, et al. Teratogenic determinants of first- trimester exposure to antiepileptic medications. Journal of Population Therapeutics and Clinical Pharmacology. 2011;18:e89-e98.

Brownell MD, Hanlon-Dearman ACB, MacWilliam LR, Chudley AE, Roos NP, Yallop LP, et al. Use of health, education, and social services by individuals with fetal alcohol spectrum disorder. Journal of Population Therapeutics and Clinical Pharmacology. 2013;20:e95-106.

Brugha TS, Wheatley S, Taub NA, Culverwell A, Friedman T, Kirwan P, et al. Pragmatic randomized trial of antenatal intervention to prevent post-natal depression by reducing psychosocial risk factors. Psychological Medicine. 2000;30:1273-81.

Brugha TS, Morrell CJ, Slade P, Walters SJ. Universal prevention of depression in women postnatally: cluster randomized trial evidence in primary care. Psychological Medicine. 2011;41:739-48.

Buist A, Gotman N, Yonkers KA. Generalized anxiety disorder: course and risk factors in pregnancy. Journal of Affective Disorders. 2011;131:277-83.

Bulik CM, Von Holle A, Siega-Riz AM, Torgersen L, Lie KK, Hamer RM, et al. Birth outcomes in women with eating disorders in the Norwegian Mother and Child cohort study (MoBa). International Journal of Eating Disorders. 2009;42:9-18.

Bunevicius A, Kusminskas L, Pop VJ, Pedersen CA, Bunevicius R. Screening for antenatal depression with the Edinburgh Depression Scale. Journal of Psychosomatic Obstetrics and Gynecology. 2009;30:238-43.

Burja S, Rakovec-Felser Z, Treiber M, Hajdinjak D, Gajsek-Marchetti M. The frequency of neonatal morbidity after exposure to antiepileptic drugs in utero: a retrospective population-based study. Wiener Klinische Wochenschrift. 2006;118:12-6.

Burns A, O'Mahen H, Baxter H, Bennert K, Wiles N, Ramchandani P. A pilot randomised controlled trial of cognitive behavioural therapy for antenatal depression. BMC Psychiatry. 2013;13:33.

Buttner MM, Mott SL, Pearlstein T, Stuart S, Zlotnick C, O'Hara MW. Examination of premenstrual symptoms as a risk factor for depression in postpartum women. Archives of Women's Mental Health. 2013;16:219-25.

Butz AM, Lears MK, O'Neil S, Lukk P. Home intervention for in utero drug-exposed infants. Public Health Nursing. 1998;15:307-18.

Butz AM, Pulsifer M, Marano N, Belcher H, Lears MK, Royall R. Effectiveness of a home intervention for perceived child behavioral problems and parenting stress in children with in utero drug exposure. Archives of Pediatrics and Adolescent Medicine. 2001;155:1029-37.

Cacciatore J, Rådestad I, Frøen F. Effects of contact with stillborn babies on maternal anxiety and depression. Birth. 2008;35:313-20.

Caldera D, Burrell L, Rodriguez K, Crowne SS, Rohde C, Duggan A. Impact of a statewide home visiting program on parenting and on child health and development. Child Abuse and Neglect. 2007;31:829-52.

Calderon-Margalit R, Qiu C, Ornoy A, Siscovick DS, Williams MA. Risk of preterm delivery and other adverse perinatal outcomes in relation to maternal use of psychotropic medications during pregnancy. American Journal of Obstetrics and Gynecology. 2009;201:579(e1-8).

Campbell S, Norris S, Standfield L, Suebwongpat A. Screening for postnatal depression within the Well Child Tamariki Ora Framework. Health Services Assessment Collaboration; 2008.

Canger R, Battino D, Canevini MP, Fumarola C, Guidolin L, Vignoli A, et al. Malformations in offspring of women with epilepsy: a prospective study. Epilepsia. 1999;40:1231-36.

Carpiniello B, Pariante CM, Serri F, Costa G, Carta MG. Validation of the Edinburgh postnatal depression scale in Italy. Journal of Psychosomatic Obstetrics and Gynaecology. 1997;18:280-5.

Carrol M, Chang G, Behr H, Clinton B, Kosten T. Improving Treatment Outcome in Pregnant, MEthadone-Maintained WOmen. American Journal on Addictions. 1995;4:56-9.

Casper RC, Fleisher BE, Lee-Ancajas JC, Gilles A, Gaylor E, DeBattista A, et al. Follow-up of children of depressed mothers exposed or not exposed to antidepressant drugs during pregnancy. Journal of Pediatrics. 2003;142:402-8.

Cassina M, Dilaghi A, Di Gianantonio E, Cesari E, De Santis M, Mannaioni G, et al. Pregnancy outcome in women exposed to antiepileptic drugs: teratogenic role of maternal epilepsy and its pharmacologic treatment. Reproductive Toxicology. 2013;39:50-7.

Chambers CD, Johnson KA, Dick LM, Felix RJ, Jones KL. Birth outcomes in pregnant women taking fluoxetine. New England Journal of Medicine. 1996;335:1010-15.

Chambers C, Hernandez-Diaz S, Mitchell AA. Selective serotonin-reuptake inhibitors and risk of persistent pulmonary hypertension of the newborn: comment reply. The New England Journal of Medicine. 2006;354:2189-90.Chang G, Wilkins-Haug L, Berman S, Goetz MA. Brief intervention for alcohol use in pregnancy: randomized trial. Addiction. 1999;94:1499-1508.

Chang G, Goetz MA, Wilkins-Haug L, Berman S. A brief intervention for prenatal alcohol use: an in-depth look. Journal of Substance Abuse Treatment. 2000;18:365-9.

Chang G, McNamara TK, Orav J, Koby D, Lavigne A, Ludman B, et al. Brief intervention for prenatal alcohol use: a randomized trial. Obstetrics and Gynecology. 2005;105:991-8.

Chang G, McNamara TK, Orav EJ, Wilkins-Haug L. Brief intervention for prenatal alcohol use: the role of drinking goal selection. Journal of Substance Abuse Treatment. 2006;31:419-24.

Charlton RA, Weil JG, Cunnington MC, Ray S, De Vries CS. Comparing the general practice research database and the UK epilepsy and pregnancy register as tools for postmarketing teratogen surveillance: anticonvulsants and the risk of major congenital malformations. Drug Safety. 2011;34:157-71.

Chaudron LH, Szilagyi PG, Tang W, Anson E, Talbot NL, Wadkins HI, Tu X, Wisner KL. Accuracy of depression screening tools for identifying postpartum depression among urban mothers. Pediatrics. 2010;125:e609-17.

Chen C-H, Tseng Y-F, Chou F-H, Wang S-Y. Effects of support group intervention in postnatally distressed women. A controlled study in Taiwan. Journal of Psychosomatic Research. 2000;49:395-99.

Chew-Graham CA, Sharp D, Chamberlain E, Folkes L, Turner KM. Disclosure of symptoms of postnatal depression, the perspectives of health professionals and women: a qualitative study. BMC Family Practice. 2009;10:7.

Chibanda D, Mangezi W, Tshimanga M, Woelk G, Rusakaniko P, Stranix-Chibanda L, et al. Validation of the Edinburgh Postnatal Depression Scale among women in a high HIV prevalence area in urban Zimbabwe. Archives of Women's Mental Health. 2010;13:201-6.

Chisolm MS, Fitzsimons H, Leoutsakos JM, Acquavita SP, Heil SH, Wilson-Murphy M, et al. A comparison of cigarette smoking profiles in opioid-dependent pregnant patients receiving methadone or buprenorphine. Nicotine Tobacco Research. 2013;15:1297-304.

Cho HJ, Kwon JH, Lee JJ. Antenatal cognitive-behavioral therapy for prevention of postpartum depression: a pilot study. Yonsei Medical Journal. 2008;49:553-562.

Christensen J, Grnoborg TK, Sroensen MJ, Schendel D, Parner ET, Pedersen LH, et al. Prenatal valproate exposure and risk of autism spectrum disorders and childhood autism. Journal of the American Medical Association. 2013;309:1696-703.

Chung KF, Yeung WF, Zhang ZJ, Yung KP, Man SC, Lee CP et al. Randomized non-invasive sham-controlled pilot trial of electroacupuncture for postpartum depression. Journal of Affective Disorders. 2012;142:115-21.

Clarke PJ. Validation of two postpartum depression screening scales with a sample of First Nations and Metis women. The Canadian Journal of Nursing Research. 2008;40:113-25.

Cochrane Collaboration. Review Manager (RevMan) [Computer program]. Version 5.2. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration. 2012.

Collins CH, Zimmerman C, Howard LM. Refugee, asylum seeker, immigrant women and postnatal depression: rates and risk factors. Archives of Women's Mental Health. 2011;14:3-11.

Conroy S, Pariante CM, Marks MN, Davies HA, Farrelly S, Schacht R, et al. Maternal psychopathology and infant development at 18 months: the impact of maternal personality disorder and depression. Journal of the American Academy of Child and Adolescent Psychiatry. 2012;51:51-61.

Cooke S, Smith I, Turl E, Arnold E, Msetfi RM. Parent perspectives of clinical psychology access when experiencing distress. Community Practitioner. 2012;85:34-7.

Cooper PJ, Murray L. Course and recurrence of postnatal depression. Evidence for the specificity of the diagnostic concept. The British Journal of Psychiatry. 1995;166:191-5.

Cooper P, Murray L, Hooper R, West A. The development and validation of a predictive index for postpartum depression. Psychological Medicine.1996;26:627-34.

Cooper PJ, Murray L, Wilson A, Romaniuk H. Controlled trial of the short- and long-term effect of psychological treatment of post-partum depression. I. Impact on maternal mood. British Journal of Psychiatry. 2003;182:412-9.

Cooper PJ, Tomlinson M, Swartz L, Landman M, Molteno C, Stein A, et al. Improving quality of mother-infant relationship and infant attachment in socioeconomically deprived community in South Africa: randomised controlled trial. BMJ. 2009;338:b974.

Correa-Villaseñor A, Ferencz C, Neill CA, Wilson PD, Boughman JA. Ebstein's malformation of the tricuspid valve: genetic and environmental factors. The Baltimore-Washington Infant Study Group. Teratology. 1994;50:137-47.

Costei AM, Kozer E, Ho T, Ito S, Koren G. Perinatal outcome following third trimester exposure to paroxetine. Archives of Pediatrics and Adolescent Medicine. 2002;156:1129-32.

Cox JL, Holden JM, Sagovsky R. Detection of postnatal depression: development of the 10-item Edinburgh Postnatal Depression scale. British Journal of Psychiatry. 1987;150:782-6.

Cox JL, Murray D, Chapman G. A controlled study of the onset, duration and prevalence of postnatal depression. The British Journal of Psychiatry. 1993;163:27-31.

Coyle MG, Salisbury AL, Lester BM, Jones HE, Lin H, Graf-Rohrmeister K, et al. Neonatal neurobehavior effects following buprenorphine versus methadone exposure. Addiction. 2012;107(Suppl. 1):63-73.

Craig MC. Should psychiatrists be prescribing oestrogen therapy to their female patients? British Journal of Psychiatry. 2013;202:9-13.

Crawley R, Lomax S, Ayers S. Recovering from stillbirth: the effects of making and sharing memories on maternal mental health. Journal of Reproductive and Infant Psychology. 2013;31:195-207.

Croen LA, Grether JK, Yoshida CK, Odouli R, Hendrick V. Antidepressant use during pregnancy and childhood autism spectrum disorders. Archives of General Psychiatry. 1104;68:1104-12.

Cunnington MC, Weil JG, Messenheimer JA, Ferber S, Yerby M, Tennis P. Final results from 18 years of the International Lamotrigine Pregnancy Registry. Neurology. 2011;76:1817-23.

Curtis L. Unit Costs of Health and Social Care 2013. Canterbury: University of Kent; 2013.

Czarnocka J, Slade P. Prevalence and predictors of post-traumatic stress symptoms following childbirth. British Journal of Clinical Psychology. 2000;39:35-51.

Czeizel A. Lack of evidence of teratogenicity of benzodiazepine drugs in Hungary. Reproductive toxicology. 1987;1:183-88.

Czeizel A, Racz J. Evaluation of drug intake during pregnancy in the Hungarian Case-Control Surveillance of Congenital Anomalies. Teratology. 1990;42:505-12.

Czeizel AE, Szegal BA, Joffe JM, Racz J. The effect of diazepam and promethazine treatment during pregnancy on the somatic development of human offspring. Neurotoxicology and Teratology. 1999;21:157-67.

Czeizel AE, Eros E, Rockenbauer M, Sorensen HT, Olsen J. Short-term oral diazepam treatment during pregnancy: a population-based teratological case-control study. Clinical Drug Investigation. 2003;23:451-62.

Czeizel AE, Rockenbauer M, Sorensen HT, Olsen J. A population-based case-control study of oral chlordiazepoxide use during pregnancy and risk of congenital abnormalities. Neurotoxicology & Teratology. 2004;26:593-8.

Dagher RK, McGovern PM, Dowd BE, Gjerdingen DK. Postpartum depression and health services expenditures among employed women. Journal of Occupational and Environmental Medicine. 2012;54:210-15.

Dakof GA, Quille TJ, Tejeda MJ, Alberga LR, Bandstra E, Szapocznik J. Enrolling and retaining mothers of substance-exposed infants in drug abuse treatment. Journal of Consulting and Clinical Psychology. 2003;71:764-72.

Daley A, Winter H, Grimmett C, McGuinness M, McManus R, MacArthur C. Feasibility of an exercise intervention for women with postnatal depression: A pilot randomised controlled trial." British Journal of General Practice. 2008;58:178-183.

Daley AJ, Blamey RV, Jolly K, Roalfe AK, Turner KM, Coleman S et al. A pragmatic randomised controlled trial to evaluate the effectiveness of exercise as a treatment for postnatal depression: the PAM-PeRS trial. (In press).

Daniels JL, Forssen U, Hultman CM, Cnattinqius S, Savitz DA, Feychting M et al. Parental psychiatric disorders associated with autism spectrum disorders in the offspring. Pediatrics. 2008;121:e1357-62.

Davey SJ, Dziurawiec S, O'Brien-Malone A. Men's voices: postnatal depression from the perspective of male partners. Qualitative Health Research. 2006:16:206-20.

Davis RL, Rubanowice D, McPhillips H, Raebel MA, Andrade SE, Smith D, et al. Risks of congenital malformations and perinatal events among infants exposed to antidepressant medications during pregnancy. Pharmacoepidemiology and Drug Safety. 2007;16:1086-94.

De Jonge A. Support for teenage mothers: a qualitative study into the views of women about the support they received as teenage mothers. Journal of Advanced Nursing. 2001;36:49-57.

De Wolff MS, Van Ijzendoorn MH. Sensitivity and attachment: A meta-analysis of parental antecedants of infant attachment. Child Dev.1997;68:571-91.

Deeks JJ. Issues in the selection of a summary statistic for meta-analysis of clinical trials with binary outcomes. *Statistics in Medicine*. 2002;*21*:1575−1600.

Dennis C-L. The effect of peer support on postpartum depression: a pilot randomized controlled trial. Canadian Journal of Psychiatry. 2003;48:115-124.

Dennis C-L. Postpartum depression peer support: maternal perceptions from a randomized controlled trial. International Journal of Nursing Studies. 2010;47:560-8.

Dennis C-L, Hodnett E, Reisman HM, Kenton L, Weston J, Zupancic J, et al. Effect of peer support on prevention of postnatal depression among high risk women: multisite randomised controlled trial. BMJ. 2009;338:a3064.

Dennis C-L, Dowswell T. Psychosocial and psychological interventions for preventing postpartum depression (Review). Cochrane Database of Systematic Reviews. 2013;2:Art. No.:CD001134.mdoi: 10.1002/14651858.CD001134.pub3.

Department of Health. National Service Framework for Mental Health: Modern Standards and Service Models. London: Department of Health. 1999a. Available at: http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_4009598 (accessed 16/6/2014)

Department of Health. Effective Care Co-ordination in Mental Health Services: Modernising the Care Programme Approach – A Policy Booklet. London: Department of Health; 1999b.

Department of Health. Women's Mental Health: into the Mainstream: Strategic Development of Mental Health Care for Women. London: Department of Health; 2002.

Department of Health. Hospital Episode Statistics. London: Department of Health. 2002. Available at: http://webarchive.nationalarchives.gov.uk/+/www.dh.gov.uk/en/Publicationsandstatistics/Statistics/StatisticalWorkAreas/Statisticalworkareaneighbourhood/DH_4116710 (accessed 16 June 2014).

Department of Health. Organising and Delivering Psychological Therapies. London: Department of Health; 2004.

Department of Health. Refocusing the Care Programme Approach: Policy and Positive Practice Guidance. London: Department of Health; 2008.

Di Florio A, Forty L, Gordon-Smith K, Heron J, Jones L, Craddock N, et al. Perinatal episodes across the mood disorder spectrum. JAMA Psychiatry. 2013;70:168-75.

Diav-Citrin O, Okotore B, Lucarelli K, Koren G. Pregnancy outcome following first-trimester exposure to zopiclone: a prospective controlled cohort study. American Journal of Perinatology. 1999;16:157-60.

Diav-Citrin O, Shechtman S, Arnon J, Ornoy A. Is carbamazepine teratogenic? A prospective controlled study of 210 pregnancies. Neurology. 2001;57:321-24.

Diav-Citrin O, Shechtman S, Ornoy S, Arnon J, Schaefer C, Garbis H, et al. Safety of haloperidol and penfluridol in pregnancy: a multicenter, prospective, controlled study. Journal of Clinical Psychiatry. 2005;66:317-22.

Diav-Citrin O, Shechtman S, Weinbaum D, Wajnberg R, Avgil M, Di Gianantonio E, et al. Paroxetine and fluoxetine in pregnancy: A prospective, multicentre, controlled, observational study. British Journal of Clinical Pharmacology. 2008a:66;695-705

Diav-Citrin O, Shechtman S, Bar-Oz B, Cantrell D, Arnon J, Ornoy A. Pregnancy outcome after in utero exposure to valproate evidence of dose relationship in teratogenic effect. CNS Drugs. 2008b;22:325-34.

Diego MA, Field T, Hernandez-Reif M, Cullen C, Schanberg S, Kuhn C. Prepartum, postpartum, and chronic depression effects on newborns. Psychiatry. 2004;67:63-80.

Dolk H, Jentink J, Loane M, Morris J, de Jong-van den Berg LT. Does lamotrigine use in pregnancy increase orofacial cleft risk relative to other malformations? Neurology. 2008;71:714-22.

Dolman C, Jones I, Howard LM. A systematic review and meta-synthesis of the experience of motherhood in women with severe mental illness. Archives of Women's Mental Health. 2013;16:173-96.

Duggan AK, Caldera D, Rodriguez K, Burrell L, Rohde C, Crowne SS. Impact of a statewide home visiting program to prevent child abuse. Child Abuse and Neglect. 2007;31:829-52.

Dugravier R, Tubach F, Saias T, Guedeney N, Pasquet B, Purper-Ouakil D, et al. Impact of a manualized multifocal perinatal home-visiting program using psychologists on postnatal depression: the CAPEDP randomized controlled trial. PLoS ONE. 2013;8:e72216.

Dukhovny D, Dennis CL, Hodnett E, Weston J, Stewart DE, Mao W, et al. Prospective economic evaluation of a peer support intervention for prevention of postpartum depression among high risk women. American Journal of Perinatology. 2013;30:631-42.

Easter A, Bye A, Taborelli E, Corfield F, Schmidt U, Treasure J, et al. Recognising the symptoms: how common are eating disorders in pregnancy? European Eating Disorders Review. 2013;21:340-4.

Easterbrooks MA, Bartlett JD, Raskin M, Goldberg J, Contreras MM, Kotake C. Limiting home visiting effects: maternal depression as a moderator of child maltreatment. Pediatrics. 2013;132:S126-S133.

Eberhard-Gran M, Eskild A, Tambs K, Schei B, Opjordsmoen S. The Edinburgh Postnatal Depression Scale: validation in a Norwegian community sample. Nordic Journal of Psychiatry. 2001;55:113-7.

Eccles M, Freemantle N, Mason J. North of England evidence based guideline development project: methods of developing guidelines for efficient drug use in primary care. British Medical Journal. 1998;316:1232−35.

Edge D. Perinatal depression and Black Caribbean women: lessons for primary care. Primary Health Care. 2007;17:32-5.

Edge D. 'We don't see Black women here': an exploration of the absence of Black Caribbean women from clinical and epidemiological data on perinatal depression in the UK. Midwifery. 2008;24:379-389.

Edge D. 'It's leaflet, leaflet, leaflet then, "see you later"': black Caribbean women's perceptions of perinatal mental health care. British Journal of General Practice. 2011;61:256-262.

Edge D, Rogers A. Dealing with it: Black Caribbean women's response to adversity and psychological distress associated with pregnancy, childbirth, and early motherhood. Social Science and Medicine. 2005;61:15-25.

Edoka IP, Petrou S, Ramchandani PG. Healthcare costs of paternal depression in the postnatal period. Journal of Affective Disorders. 2011;133:356-60.

Edwards E, Timmons S. A qualitative study of stigma among women suffering postnatal illness. Journal of Mental Health. 2005;14:471-81.

Egan G. The Skilled Helper: a Systematic Approach to Effective Helping. Pacific Grove, CA: Brooks/Cole; 1990.

Einarson RN, Pistelli A, DeSantis MD, Malm H, Paulus W, Panchaud A. Evaluation of the Risk of Congenital cardiovascular Defects Associated With Use of Paroxetine During Pregnancy. American Journal of Psychiatry. 2008;165:749-752

Einarson A, Boskovic R. Use and safety of antipsychotic drugs during pregnancy. Journal of Psychiatric Practice. 2009;15:183-92.

Ekeroma AJ, Ikenasio-Thorpe B, Weeks S, Kokaua J, Puniani K, Stone P et al. Validation of the Edinburgh Postnatal Depression Scale (EPDS) as a screening tool for postnatal depression in Samoan and Tongan women living in New Zealand. New Zealand Medical Journal. 2012;125:41-50.

El Marroun H, White TJH, Van der Knaap NJF, Homberg JR, Fernandez G, Schoemaker NK et al. Prenatal exposure to selective serotonin reuptake inhibitors and autistic symptoms in young children: population-based study of young children. The British Journal of Psychiatry. 2014;205:95-102.

El-Mohandes AAE, Kiely M, Joseph JG, Subramanian S, Johnson AA, Blake SM, et al. An intervention to improve postpartum outcomes in African-American mothers: a randomized controlled trial. Obstetrics and Gynecology. 2008;112: 611-20.

Elk R, Mangus L, Rhoades H, Andres R, Grabowski J. Cessation of cocaine use during pregnancy: effects of contingency management interventions on maintaining abstinence and complying with prenatal care. Addictive Behaviors. 1998;23:57-64.

Elliott SA, Leverton TJ, Sanjack M, Turner H, Cowmeadow P, Hopkins J, et al. Promoting mental health after childbirth: a controlled trial of primary prevention of postpartum depression. British Journal of Clinical Psychology. 2000;39:223–41.

Eriksson K, Viinikainen K, Monkkonen A, Aikia M, Nieminen P, Heinonen S, et al. Children exposed to valproate in utero--population based evaluation of risks and confounding factors for long-term neurocognitive development. Epilepsy Research. 2005;65:189-200.

Ernst CC, Grant TM, Streissguth AP, Sampson P. Intervention with high-risk alcohol and drug abusing mothers. II. Three-year findings from the Seattle model of paraprofessional advocacy. Journal of Community Psychology. 1999;27:19-38.

Eros E, Czeizel AE, Rockenbauer M, Sorensen HT, Olsen J. A population-based case-control teratologic study of nitrazepam, medazepam, tofisopam, alprazolum and clonazepam treatment during pregnancy. European Journal of Obstetrics Gynecology and Reproductive Biology. 2002;101:147-54.

Ethen MK, Ramadhani TA, Scheuerle AE, Canfield MA, Wyszynski DF, Druschel CM, et al. National Birth Defects Prevention Study. Alcohol consumption by women before and during pregnancy. Maternal Child Health Journal. 2009;13:274-85.

Evans J, Heron J, Francomb H, Oke S, Golding J. Cohort study of depressed mood during pregnancy and after childbirth. British Medical Journal. 2001;323:257-60.

Fahy TA, Morrison JJ. The clinical significance of eating disorders in obstetrics. British Journal of Obstetrics and Gynaecology. 1993;100:708-10.

Feeley N, Zelkowitz P, Shrier I, Stremler R, Westreich R, Dunkley D, et al. Follow-up of the cues and care trial: mother and infant outcomes at 6 months. Journal of Early Intervention. 2012;34:65-81.

Felice E, Saliba J, Grech V, Cox J. Validation of the Maltese version of the Edinburgh Postnatal Depression Scale. Archives of Women's Mental Health. 2006;9:75-80.

Fernandes MC, Srinivasan K, Stein AL, Menezes G, Sumithra R, Ramchandani PG. Assessing prenatal depression in the rural developing world: a comparison of two screening measures. Archives of Womens Mental Health. 2011;14:209-16.

Ferreira E, Carceller AM, Agogue C, Martin BZ, St-Andre M, Francoeur D, et al. Effects of selective serotonin reuptake inhibitors and venlafaxine during pregnancy in term and preterm neonates. Pediatrics. 2007;119:52-9.

Field T. Postpartum depression effects on early interactions, parenting and safety practices: a review. Infant Behaviour Development. 2010;33:1-6.

Field T, Diego M, Delgado J, Medina L. Peer support and interpersonal psychotherapy groups experienced decreased prentatal depression, anxiety and cortisol. Early Human Development. 2013a;89:621-624.

Field T, Diego M, Delgado J, Medina L. Tai chi/yoga reduces prenatal depression, anxiety and sleep disturbances. Complementry Therapy Clinical Practice. 2013b;19:6-10.

Fischer G, Jagsch R, Eder H, Gombas W, Etzerdorfer P, Schmidl-Mohl K, et al. Comparison of methadone and slow-release morphine maintenance in pregnant addicts. Addiction. 1999;94:231-9.

Fischer G, Ortner R, Rohrmeister K, Jagsch R, Baewert A, Langer M, et al. Methadone versus buprenorphine in pregnant addicts: a double-blind, double dummy comparison study. Addiction. 2006;101:275-81.

Fleming MF, Lund MR, Wilton G, Landry M, Scheets D. The Healthy Moms study: the efficacy of brief alcohol intervention in postpartum women. Alcoholism: Clinical and Experimental Research. 2008;32:1600-06.

Flynn HA, Sexton M, Ratliff S, Porter K, Zivin K. Comparative performance of the Edinburgh Postnatal Depression Scale and the Patient Health Questionnaire-9 in pregnant and postpartum women seeking psychiatric services. Psychiatry Research. 2011;187:130-4.

Flynn SM, Shaw JJ, Abel KM. Homicide of infants: a cross-sectional study. Journal of Clinical Psychiatry. 2007;68:1501-9.

Fonager K, Larsen H, Pedersen L, Sorensen HT. Birth outcomes in women exposed to anticonvulsant drugs. Acta Neurologica Scandinavica. 2000;101:289-94.

Forsberg L, Wide K, Kallen B. School performance at age 16 in children exposed to antiepileptic drugs in utero- A population-based study. Epilepsia. 2011;52:364-9.

Fraser JA, Armstrong KL, Morris JP, Dadds MR. Home visiting intervention for vulnerable families with newborns: follow-up results of a randomized controlled trial. Child Abuse & Neglect. 2000;24:1399-429.

Freeman MP, Davis M, Sinha P, Wisner KL, Hibbeln JR, Gelenberg AJ. Omega-3 fatty acids and supportive psychotherapy for perinatal depression: A randomized placebo-controlled study. Journal of Affective Disorders. 2008;110:142-8.

Furuta M, Sandall J, Bick D. A systematic review of the relationship between severe maternal morbidity and post-traumatic stress disorder. BMC Pregnancy and Childbirth. 2012;12:5.

Gaalema DE, Scott TL, Heil SH, Coyle MG, Kaltenbach K, Badger GJ, et al. Differences in the profile of neonatal abstinence syndrome signs in methadone- versus buprenorphine-exposed neonates. Addiction. 2012;107 (Suppl. 1):53-62.

Gagnon AJ. Individual or group antenatal education for childbirth/parenthood. Cochrane Database of Systematic Reviews. 2000;4:Art. No.: CD002869. DOI: 10.1002/14651858.CD002869.pub2

Gaily E, Kantola-Sorsa E, Hiilesmaa V, Isoaho M, Matila R, Kotila M, et al. Normal intelligence in children with prenatal exposure to carbamazepine. Neurology. 2004;62:28-32.

Galbally M, Lewis A, Lum J, Buist A. Serotonin discontinuation syndrome following in utero exposure to antidepressant medication: prospective controlled study. Australian and New Zealand Journal of Psychiatry. 2009;43:846-54.

Gallivan S, Utley M, Treasure T, Valencia O. Booked inpatient admissions and capacity: mathematical modelling study. British Medical Journal. 2002;324:280-2.

Gamble J, Creedy D, Moyle W, Webster J, McAllister M, Dickson P. Effectiveness of a counseling intervention after a traumatic childbirth: a randomized controlled trial. Birth. 2005;32:11-9.

Gao L-L, Chan SW-C, Li X, Chen S, Hao Y. Evaluation of an interpersonal-psychotherapy-oriented childbirth education programme for Chinese first-time childbearing women: a randomised controlled trial. International Journal of Nursing Studies. 2010;47:1208-16.

Gao L-L, Chan SW-C, Sun K. Effects of an interpersonal-psychotherapy-oriented childbirth education programme for Chinese first-time childbearing women at 3-month follow up: randomised controlled trial. International Journal of Nursing Studies. 2012;49:274-81.

Garcia-Esteve L, Ascaso C, Ojuel J, Navarro P. Validation of the Edinburgh Postnatal Depression Scale (EPDS) in Spanish mothers. Journal of Affective Disorders. 2003;75:71-6.

Gausia K, Fisher C, AlgincS, OosthuizendJ. Validation of the Bangla version of the Edinburgh Postnatal Depression Scale for a Bangladeshi sample. Journal of Reproductive and Infant Psychology. 2007;25:308-15.

Gavin NI, Gaynes BN, Lohr KN, Meltzer-Brody S, Gartlehner G, et al. Perinatal depression: a systematic review of prevalence and incidence. Obstetrics and Gynecology. 2005;106;1071-83.

Ghubash R, Abou-Saleh MT, Daradkeh TK.The validity of the Arabic Edinburgh postnatal depression scale.Social Psychiatry and Psychiatric Epidemiology. 1997;32:474-6.

Gibson J, McKenzie-McHarg K, Shakespeare J, Price J, Gray R. A systematic review of studies validating the Edinburgh Postnatal Depression Scale in antepartum and postpartum women. Acta Psychiatrica Scandinavica. 2009;119:350-64

Gilbody S, Richards D, Brealey S, Hewitt C. Screening for depression in medical settings with the Patient Health Questionnaire (PHQ): a diagnostic meta-analysis. Journal of General Internal Medicine. 2007;22:1596-1602.

Gilinsky A, Swanson V, Power K. Interventions delivered during antenatal care to reduce alcohol consumption during pregnancy: a systematic review. Addiction Research and Theory. 2011;19:235-50.

Gjerdincjen D, Crow S, McGovern P, Miner M, Center B. Postpartum depression screening at well-child visits: Validity of a 2-question screen and the PHQ-9. Annals of Family Medicine. 2009;7:63-70.

Goldberg D, Privett M, Ustun B, Simon G, Linden M. The effects of detection and treatment on the outcome of major depression in primary care: a naturalistic study in 15 cities. The British Journal of General Practice. 1998:48;1840-4.

Goldberg DP, Bridges K. Somatic presentations of psychiatric illness in primary care setting. Journal of Psychosomatic Research. 1988;32:137-44.

Goldberg DP, Huxley P. Common Mental Disorders: A Bio-Social Model. London: Tavistock/Routledge; 1992.

Goldberg DP, Jenkins L, Millar T, Faragher EB.The ability of trainee general practitioners to identify psychological distress among their patients. Psychological Medicine. 1993:23;185-93.

Goldenberg RL, Culhane JF, Iams JD, Romero R. Preterm birth: what can be done? The Lancet. 2008;371:75-84.

Goodman JH. Postpartum depression beyond the early postpartum period. Journal of Obstetric, Gynecologic and Neonatal Nursing. 2004;33:410-20.

Goodwin N, Perri6, Peck E, Freeman T, Posaner R. Managing Across Diverse Networks of Care: Lessons from Other Sectors. Policy Report to the NHS SDO R & D Programme, Health Services Management Centre, University of Birmingham. 2004

Gorman L. Prevention of postpartum difficulties in a high risk sample [dissertation]. Iowa City, IA: University of Iowa; 1997.

Grant TM, Ernst CC, Streissguth AP. An intervention with high-risk mothers who abuse alcohol and drugs: the Seattle Advocacy Model. American Journal of Public Health. 1996a;86:1816-7.

Grant TM, Ernst CC, Streissguth AP, Phipps P, Gendler B. When case management isn't enough: a model of paraprofessional advocacy for drug- and alcohol-abusingmothers. Journal of Case Management. 1996b;5:3-11.

Grant TM, Ernst CC, Streissguth A, Stark K. Preventing alcohol and drug exposed births in Washington state: intervention findings from three parent-child assistance program site. American Journal of Drug and Alcohol Abuse. 2005;31:471-90.

Gravensteen IK, Helgadóttir LB, Jacobsen E-M, Rådestad I, Sandset PM, et al. Women's experiences in relation to stillbirth and risk factors for long-term post-traumatic stress symptoms: a retrospective study. BMJ Open. 2013;3:e003323.

Gray R, Mukherjee RAS, Rutter M. Alcohol consumption during pregnancy and its effects on neurodevelopment: what is known and what remains uncertain. Addiction. 2009;104:1270-3.

Gray R. Life stage: Pre-conception and Pregnancy. Annual Report of the Chief Medical Officer 2012, Our Children Deserve Better: Prevention Pays. London: Department of Health; 2013. Available from: http://socialwelfare.bl.uk/subject-areas/services-client-groups/children-young-people/departmentofhealth/15526533571_2901304_CMO_All.pdf

Gregoire AJ, Kumar R, Everitt B, Henderson AF, Studd JWW. Transdermal oestrogen for treatment of severe postnatal depression. Lancet. 1996;347:930-933

Grekin R, O'Hara MW. Prevalence and risk factors of postpartum posttraumatic stress disorder: A meta-analysis. Clinical Psychology Review. 2014;34:389-401.

Grigoriadis S, VonderPorten EH, Mamisashvili L, Roerecke M, Rehm J, Dennis CL, et al. Antidepressant exposure during pregnancy and congenital malformations: is there an association? A systematic review and meta-analysis of the best evidence. Journal of Clinical Psychiatry. 2013a:74;e293-e308.

Grigoriadis S, VonderPorten EH, Mamisashvili L, Tomlinson G, Dennis CL, Koren G et al. 2013b. Prenatal exposure to antidepressants and persistent pulmonary hypertension of the newborn: systematic review and meta-analysis. British Medical Journal. 2013b:348;f6932.

Grigoriadis S, VonderPorten EH, Mamisashvili L, Eady A, Tomlinson G, Dennis CL et al. The effect of prenatal antidepressant exposure on neonatal adaptation: a

systematic review and meta-analysis. Journal of Clinical Psychiatry. 2013c:74;e309-e320.

Grote NK, Swartz HA, Geibel SL, Zuckoff A, Houck PR, Frank E. A randomized controlled trial of culturally relevant, brief interpersonal psychotherapy for perinatal depression. Psychiatric Services. 2009;60:313-21.

Grote NK, Bridge JA, Gavin AR, Melville JL, Iyengar S, Katon WJ. A metanalysis of depression during pregnancy and the risk of preterm birth, low birth weight and intrauterine growth restriction. Archives of General Psychiatry. 2010;67:1012-24.

Guardino CM, Schetter CD, Bower JE, Lu MC, Smalley SL. Randomised controlled pilot trial of mindfulness training for stress reduction during pregnancy. Psychology and Health. 2014;29:334-49.

Guedeney A, Wendland J, Dugravier R, Saias T, Tubach F, Welniarz B, et al. Impact of a randomized home-visiting trial on infant social withdrawal in the CAPEDP prevention study. Infant Mental Health Journal. 2013;34:594-601.

Guedeney N, Fermanian J. Validation study of the French version of the Edinburgh Postnatal Depression Scale (EPDS): new results about use and psychometric properties. European Psychiatry. 1998;13:83-9.

Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. Journal of Clinical Epidemiology. 2011;64:383-94.

Haddad-Rodrigues M, Nakano AMS, Stefanello J, Silveira RCCP. Acupuncture for anxiety in lactating mothers with preterm infants: a randomized controlled trial. Evidence-Based Complementary and Alternative Medicine. 2013;2013:169-84.

Habermann F, Fritzsche J, Fuhlbruck F, Wacker E, Allignol A, Weber-Schoendorfer C, et al. Atypical antipsychotic drugs and pregnancy outcome: a prospective, cohort study. Journal of Clinical Psychopharmacology. 2013;33:453-62.

Hagan R, Evans SF, Pope S. Preventing postnatal depression in mothers of very preterm infants: a randomised controlled trial. BJOG. 2004;111:641-7.

Hall P. Mothers' experiences of postnatal depression: an interpretative phenomenological analysis. Community Practitioner. 2006;79:256-60.

Handmaker N, Miller W, Manicke M. Findings of a pilot study of motivational interviewing with pregnant drinkers. Journal of Studies on Alcohol. 1999;60:285-7.

Hanley J, Long B. A study of Welsh mothers' experiences of postnatal depression. Midwifery. 2006;22:147-57.

Hanlon C, Medhin G, Alem A, Araya M, Abdulahi A, Hughes M. Detecting perinatal common mental disorders in Ethiopia: validation of the self-reporting questionnaire and Edinburgh Postnatal Depression Scale. Journal of Affective Disorders. 2008;108:251-62.

Hantsoo L, Ward-O'Brien D, Czarkowski KA, Gueorguieva, R. Price, LH, Epperson CN. A randomised, placebo-controlled, double-blind trial of sertraline for postpatrum depression. Psychopharmacology. 2014;231:939-48

Hanusa BH, Scholle SH, Haskett RF, Spadaro K, Wisner KL.Screening for depression in the postpartum period: a comparison of three instruments. Journal of Women's Health. 2008;17:585-96.

Harker R. Children in Care, England: Statistics. London: House of Commons Library. Standard Note SN/SG/4470; 2012.

Harlow BL, Vitonis AF, Sparen P, Cnattingius S, Joffe H, Hultman CM. Incidence of hospitalization for postpartum psychotic and bipolar episodes in women with and without prior prepregnancy or prenatal psychiatric hospitalizations. Archives of General Psychiatry. 2007;64:42-8.

Harris B, Huckle P, Thomas R, Johns S, Fung H. The use of rating scales to identify post-natal depression. British Journal of Psychiatry. 1989;154:813-7.

Harris B, Oretti R, Lazarus J, Parkes A, John R, Richards C et al. Randomised trial of thyroxine to prevent postnatal depression in thyroid-antibody-positive women. The British Journal of Psychiatry. 2002;180:327-30.

Harris T, Brown GW, Hamilton V, Hodson S, Craig TKJ. The Newpin Antenatal and Postnatal Project: a randomised controlled trial of an intervention for perinatal depression. Poster prepared for the HSR Open Day, 6 July 2006. Institute of Psychiatry, Kings College London.

Harrison-Hohner J, Coste S, Dorato V, Curet LB, McCarron D, Hatton D. Prenatal calcium supplementation and postpartum depression: an ancillary study to a randomised trial of calcium for prevention of preeclampsia. Archives of Women's Mental Health. 2001;3:141-6.

Haug N, Svikis D, DiClemente C. Motivational enhancement therapy for nicotine dependence in methadone- maintained pregnant women. Psychology of Addictive Behaviors. 2004;18:289-92.

Hayden T, Perantie DC, Nix BD, Barnes LD, Mostello DJ, Holcomb WL, et al. Treating prepartum depression to improve infant developmental outcomes: a study of diabetes in pregnancy. Journal of Clinical Psychology in Medical Settings. 2012;19:285-92.

Henderson J, Gray R, Brocklehurst P. Systematic review of effects of low–moderate prenatal alcohol exposure on pregnancy outcome. BJOG: An International Journal of Obstetrics and Gynaecology. 2007;114:243-52.

Hendrick V, Altshuler LL, Suri R. Hormonal changes in the postpartum and implications for postpartum depression. Psychosomatics. 1998:39:93-101.

Henshaw C. Mood disturbance in the early puerperium: a review. Archives of Women's Mental Health. 2003;6:33-42.

Hernandez-Diaz S, Smith CR, Shen A, Mittendorf R, Hauser WA, Yerby M, et al. Comparative safety of antiepileptic drugs during pregnancy. Neurology. 2012:78;1692-9.

Heron J, O'Connor T G, Evans J, Golding J, Glover V, ALSPAC Study Team. The course of anxiety and depression through pregnancy and the postpartum in a community sample. Journal of Affective Disorders. 2004;80:65-73.

Heron J, McGuinness M, Blackmore ER, Craddock N, Jones I. Early postpartum symptoms in puerperal psychosis. British Journal of Obstetricians and Gynaecologists. 2008;115:348-53.

Heron J, Gilbert N, Dolman C, Shah S, Beare I, Dearden S, et al. Information and support needs during recovery from postpartum psychosis. Archives of Women's Mental Health. 2012;15:155-65.Hewitt C, Gilbody S, Brealey S, Paulden M, Palmer S, Mann R, et al. Methods to identify postnatal depression in primary care: an integrated evidence synthesis and value of information analysis. Health technology assessment. 2009;13:1-145.

Hiscock H, Wake M. Randomised controlled trial of behavioural infant sleep intervention to improve infant sleep and maternal mood. British Medical Journal. 2002;324:1062-5.

Hiscock H, Bayer J, Gold L, Hampton A, Ukoumunne OC, Wake M. Improving infant sleep and maternal mental health: a cluster randomised trial. Archives of Disease Childhood. 2007;92:952-8.

Hiscock H, Bayer JK, Hampton A, Ukomunne OC, Wake M. Long-term mother and child mental health effects of a population-based infant sleep intervention: cluster-randomized, controlled trial. Pediatrics. 2008;122:e621-7.

Hobson RP1, Patrick MP, Hobson JA, Crandell L, Bronfman E, Lyons-Ruth K. How mothers with borderline personality disorder relate to their year-old infants. British Journal of Psychiatry. 2009;195:325-30.

Holbrook AM, Baxter JK, Jones HE, Heil SH, Coyle MG, Martin PR, et al. Infections and obstetric outcomes in opioid-dependent pregnant women maintained on methadone or buprenorphine. Addiction. 2012;107(Suppl. 1):83-90.

Holden JM, Sagovsky R, Cox JL. Counselling in a general practice setting: controlled study of health visitor intervention in treatment of postnatal depression. British Medical Journal. 1989;298:223-26.

Holmes LB. Looking for long-term effects from prenatal exposures to anticonvulsants. Teratology. 2001;64:175-76.

Holmes TH, Rahe RH. The Social Readjustment Rating Scale. Journal of Psychosomatic Research. 1967;11:213-18.

Holmes LB, Baldwin EJ, Smith CR. Increased frequency of isolated cleft palate in infants exposed to lamotrigine during pregnancy. Neurology. 2008;70:2152-2158.

Honey KL, Bennett P, Morgan M. A brief psycho-educational group intervention for postnatal depression. British Journal of Clinical Psychology. 2002;41:405-9.

Horowitz JA, Bell M, Trybulski J, Munro BH, Moser D, Hartz SA, et al. Promoting responsiveness between mothers with depressive symptoms and their infants. Journal of Nursing Scholarship. 2001;33:323-9.

Howard LM, Kumar C, Thornicroft G. The psychosocial characteristics of mothers with psychotic disorders. British Journal of Psychiatry. 2001;178:427-32

Howard LM, Leese M, Kumar C, Thornicroft G. The general fertility rate in women with psychotic disorders. American Journal of Psychiatry. 2002;159:991-7.

Howard LM, Shah N, Salmon M, Appleby L. Predictors of parenting difficulties in women admitted to mother and baby units. Social Psychiatry and Psychiatric Epidemiology. 2003;38:450-5.

Howard LM, Thornicroft G, Salmon M, Appleby L. Predictors of parenting outcome in women with psychotic disorders discharged from mother and baby units. Acta Psychiatrica Scandinavica 2004;110:347-55.

Howard LM. Fertility and pregnancy in women with psychotic disorders. European Journal of Obstetrics, Gynecology and Reproductive Biology. 2005;119:3-10.

Howard LM, Kirkwood G, Latinovic R. Sudden infant death syndrome and maternal depression. Journal of Clinical Psychiatry. 2007;68:1279-83.

Howard LM, Oram S, Galley H, Trevillion K, Feder G. Domestic violence and perinatal mental disorders: a systematic review and meta-analysis. PLOS Medicine. 2013. DOI: 10.1371/journal.pmed.1001452

Howell EA, Balbierz A, Wang J, Parides M, Zlotnick C, Leventhal H. Reducing postpartum depressive symptoms among black and latina mothers: a randomized controlled trial. Obstetrics and Gynecology. 2012;119:942-49.

Howell EA, Bodnar-Derens, Balbierz A, Loudon H, Mora PA, Zlotnick C, et al. An intervention to reduce postpartum depressive symptoms: a randomized controlled trial. Archives of Womens Mental Health. 2014;17:57-63.

Hughes P, Turton P, Hopper E, Evans CDH. Assessment of guidelines for good practice in psychosocial care of mothers after stillbirth: a cohort study. The Lancet. 2002;306:114-8.

Hunt K, France E, Ziebland S, Field K, Wyke S. 'My brain couldn't move from planning a birth to planning a funeral': a qualitative study of parents' experiences of decisions after ending a pregnancy for fetal abnormality. International Journal of Nursing Studies. 2009;46:1111-21.

Hvas CL, Henriksen TB, Ostergaard JR, Dam M. Epilepsy and pregnancy: effect of antiepileptic drugs and lifestyle on birthweight. British Journal of Obstetrics and Gynaecology. 2000;107:896-902.

Institute of Medicine. Crossing the Quality Chasm: A New Health System for the 21st Century. Washington DC: National Academy Press. 2001.

InterTASC Information Specialists' SubGroup. Search Filter Resource. York: Centre for Reviews and Dissemination, 2007. Available from: http://www.york.ac.uk/inst/crd/intertasc/

Jablensky AV, Morgan V, Zubrick SR, Bower C, Yellachich LA. Pregnancy, delivery, and neonatal complications in a population cohort of women with schizophrenia and major affective disorders. The American Journal of Psychiatry. 2005;162:79-91.

Jacobson SJ, Jones K, Johnson K, Ceolin L, Kaur P, Sahn D, et al. Prospective multicentre study of pregnancy outcome after lithium exposure during first trimester. Lancet. 1992;339:530-3.

Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavaghan DJ, et al. Assessing the quality of reports of randomised clinical trials: is blinding necessary? Controlled Clinical Trials. 1996;17:1-12.

Jadresic E, Araya R, Jara C. Validation of the Edinburgh postnatal depression scale (EPDS) in Chilean postpartum women. Journal of Psychosomatic Obstetrics and Gynaecology. 1995;16:187-91.

Jansson LM, Dipietro JA, Velez M, Elko A, Williams E, Milio L, et al. Fetal neurobehavioral effects of exposure to methadone or buprenorphine. Neurotoxicology and Teratology. 2011;33:240-3.

Jardri R, Pelta J, Maron M, Thomas P, Delion P, Codaccioni X et al. Predictive validation study of the Edinburgh Postnatal Depression Scale in the first week after delivery and risk analysis for postnatal depression. Journal of Affective Disorders. 2006;93:169-76.

Jentink J, Loane MA, Dolk H, Barisic I, Garne E, Morris JK, et al. Valproic acid monotherapy in pregnancy and major congenital malformations. New England Journal of Medicine. 2010;362:2185-93.

Ji S, Long Q, Newport DJ, Na H, Knight B, Zach EB, et al. Validity of depression rating scales during pregnancy and the postpartum period: impact of trimester and parity. Journal of Psychiatric Research. 2011;45:213-9.

Johnson KC, LaPrairie JL, Brennan PA, Stowe ZN, Newport DJ. Prenatal antipsychotic exposure and neuromotor performance during infancy. Archives of General Psychiatry. 2012;69:787-94.

Jones HE, Stitzer M, Svikis D. Improving treatment outcomes for pregnant drug-dependent women using low-magnitude voucher incentives. Addictive Behaviors. 2000;25:263-7.

Jones HE, Haug N, Silverman K, Stitzer M, Svikis D. The effectiveness of incentives in enhancing treatment attendance and drug abstinence in methadone-maintained pregnant women. Drug and Alcohol Dependence. 2001;61:297-306.

Jones HE, Johnson RE, Jasinski DR, O'Grady KE, Chisholm CA, Choo RE, et al. Buprenorphine versus methadone in the treatment of pregnant opioid-dependent patients: effects on the neonatal abstinence syndrome. Drug and Alcohol Dependence. 2005;79:1-10.

Jones HE, Johnson RE, O'Grady KE, Jasinski DR, Tuten M, Milio L. Dosing adjustments in postpartum patients maintained on buprenorphine or methadone. Journal of Addiction Medicine. 2008;2:103-07.

Jones HE, Kaltenbach K, Heil SH, Stine SM, Coyle MG, Arria AM, et al. Neonatal abstinence syndrome after methadone or buprenorphine exposure. New England Journal of Medicine. 2010;363:2320-31.

Jones HE, Heil SH, Baewert A, Arria AM, Kaltenbach K, Martin PR, et al. Buprenorphine treatment of opioid-dependent pregnant women: a comprehensive review. Addiction. 2012a;107 (Suppl. 1):5-27.

Jones HE, Fischer G, Heil SH, Kaltenbach K, Martin PR, Coyle MG, et al. Maternal Opioid Treatment: Human Experimental Research (MOTHER)--approach, issues and lessons learned. Addiction. 2012b;107(Suppl. 1):28-35.

Jones I, Craddock N. Familiality of the puerperal trigger in bipolar disorder: results of a family study. American Journal of Psychiatry. 2001;158:913-17.

Jones I, Craddock N. Bipolar disorder and childbirth: the importance of recognising risk. British Journal of Psychiatry. 2005;186:453-4.

Jones SC, Melville J, McDonald L. Prescribing sodium valproate to women of childbearing age. Royal College of Psychiatrist's International Congress; 24-27 June 2014; London, England.

Kaaja E, Kaaja R, Matila R, Hiilesmaa V. Enzyme-inducing antiepileptic drugs in pregnancy and the risk of bleeding in the neonate. Neurology. 2002;58:549-53.

Kaaja E, Kaaja R, Hiilesmaa V. Major malformations in offspring of women with epilepsy. Neurology. 2003;60:575-9.

Kaaya SF, Blander J, Antelman G, Cyprian F, Emmons KM, Matsumoto K, et al. Randomized controlled trial evaluating the effect of an interactive group counseling intervention for HIV-positive women on prenatal depression and disclosure of HIV status. AIDS Care: Psychological and Socio-medical Aspects of AIDS/HIV. 2013;25:854-62.

Kadir AA, Nordin R, Shaiful BI, Mohd JY, Wan Mohd RWM. Validation of the Malay version of Edinburgh Postnatal Depression Scale for postnatal women in Kelantan, Malaysia. International Medical Journal. 2005;12:105-9.

Kalinauskiene L, Cekuoliene D, Van Ijzendoorn MH, Bakermans-Kranenburg MJ, Juffer F, Kusakovskaja I. Supporting insensitive mothers: the Vilnius randomized control trial of video-feedback intervention to promote maternal sensitivity and infant attachment security. Child: care, health and development. 2009;35:613-23.

Kallen B. Neonate characteristics after maternal use of antidepressants in late pregnancy. Archives of Pediatrics & Adolescent Medicine. 2004;158:312-6.

Kallen B. Congenital malformations in infants whose mothers reported the use of folic acid in early pregnancy in Sweden. A prospective population study. Congenital Anomalies. 2007;47:119-24.

Kallen B, Tandberg A. Lithium and pregnancy. A cohort study on manic-depressive women. Acta Psychiatrica Scandinavica. 1983;68:134-9.

Kalotra C. Estimated Costs Related to the Birth of a Drug and/or Alcohol Exposed Baby. Washington, DC: Office of Justice Programs Drug Court Clearninghouse and Technical Assistance Project. U.S. Department of Justice; 2002.

Kammerer M, Marks MN, Pinard C, Taylor A, von Castelberg B, Künzli H, et al. Symptoms associated with the DSM IV diagnosis of depression in pregnancy and post partum. Archives of Women's Mental Health. 2009;12:135-41.

Kaneko S, Battino D, Andermann E, Wada K, Kan R, Takeda A, et al. Congenital malformations due to antiepileptic drugs. Epilepsy Research. 1999;33:145-58.

Kartin D, Grant TM, Streissguth AP, Sampson PD, Ernst CC. Three-year developmental outcomes in children with prenatal alcohol and drug exposure. Pediatric Physical Therapy. 2002;14:145-53.

Katon WJ, Russo JE, Melville JL, Katon JG, Gavin AR. Depression in pregnancy is associated with preexisting but not pregnancy-induced hypertension. General Hospital Psychiatry. 2012;34:9-16.

Kendell RE, Chalmers JC, Platz C. Epidemiology of puerperal psychoses. The British Journal of Psychiatry. 1987;150:662-73.

Kersten-Alvarez LE, Hosman CMH, Riksen-Walraven JM, van Doesum KTM, Hoefnagels C. Long-term effects of a home-visiting intervention for depressed mothers and their infants. Journal of child psychology and psychiatry, and allied disciplines. 2010;51:1160-70.

Kersting A, Kroker K, Schlicht S, Baust K, Wagner B. Efficacy of cognitive behavioral internet-based therapy in parents after the loss of a child during pregnancy: pilot data from a randomized controlled trial. Archives of Womens Mental Health. 2011;14:465-77.

Kersting A, Dölemeyer R, Steinig J, Walter F, Kroker K, Baust K, et al. Brief internet-based intervention reduces posttraumatic stress and prolonged grief in parents after the loss of a child during pregnancy: a randomized controlled trial. Psychotherapy and Psychosomatics. 2013;82:372-81.

Kessler D, Bennewith O, Lewis G, Sharp D. Detection of depression and anxiety in primary care: follow up study. BMJ. 2002;325:1016-17.

Kettinger LA, Nair P, Schuler ME. Exposure to environmental risk factors and parenting attitudes among substance-abusing women. American Journal of Drug and Alcohol Abuse. 2000;26:1-11.

Kieffer EC, Caldwell CH, Welmerink DB, Welch KB, Sinco BR, Guzmán JR. Effect of the healthy MOMs lifestyle intervention on reducing depressive symptoms among pregnant Latinas. American Journal of Community Psychology. 2013;51:76-89.

Kieler H, Artama M, Engeland A, Ericsson O, Furu K, Gissler M, et al. Selective serotonin reuptake inhibitors during pregnancy and risk of persistent pulmonary hypertension in the newborn: Population based cohort study from the five Nordic countries. British Medical Journal. 2012;344:1-9.

Killaspy H, Bebbington P, Blizard R., Johnson S, Nolan F, Pilling S, et al. The REACT study: randomised evaluation of assertive community treatment in north London. British Medical Journal. 2006;332:815-20.

King S, Griffin S, Hodges Z, Weatherly H, Asseburg C, Richardson G, et al. A Systematic Review and Economic Model of the Effectiveness and Cost-Effectiveness of Methylphenidate, Dexamfetamine and Atomoxetine for the Treatment of Attention Deficit Hyperactivity Disorder in Children and Adolescents. Health Technology Assessment. Vol. 10. Tunbridge Wells: Gray Publishing; 2006.

King-Hele SA, Abel KM, Webb RT, Mortensen PB, Appleby L, Pickles AR. Risk of sudden infant death syndrome with parental mental illness. Archives of General Psychiatry. 2007;64:1323-30.

King-Hele SA, Webb RT, Mortensen PB, Appleby L, Pickles A, Abel KM. Risk of stillbirth and neonatal death linked with maternal mental illness: a national cohort study. Archives of Disease in Childhood. Fetal and Neonatal Edition. 2009;94:F105-10.

Kini U, Lee R, Jones A, Smith S, Ramsden S, Fryer A, et al. Influence of the MTHFR genotype on the rate of malformations following exposure to antiepileptic drugs in utero. European Journal of Medical Genetics. 2007;50:411-20.

Kisely S, Gater R, Goldberg DP. Results from the Manchester Centre. In: Üstün TB, Sartorius N, eds. Mental Illness in General Health Care: an International Study. Chichester: Wiley; 1995.

Kjaer D, Horvath-Puho E, Christensen J, Vestergaard M, Czeizel AE, Sorensen HT, et al. Use of phenytoin, phenobarbital, or diazepam during pregnancy and risk of congenital abnormalities: a case-time-control study. Pharmacoepidemiology and Drug Safety. 2007;16:181-8.

Klerman GL, Weissman MM, Rounsaville BJ, Chevron ES. Interpersonal Psychotherapy of Depression. New York: Basic Books; 1984.

Knoph C, Von Holle A, Zerwas S, Torgersen L, Tambs K, Stoltenberg C, et al. Course and predictors of maternal eating disorders in the postpartum period. International Journal of Eating Disorders. 2013;46:355-368

Kornum JB, Nielsen RB, Pedersen L, Mortensen PB, Norgaard M. Use of selective serotonin-reuptake inhibitors during early pregnancy and risk of congenital malformations: updated analysis. Clinical Epidemiology. 2010;2:29-36.

Kozinszky Z, Dudas RB, Devosa I, Csatordai S, Tóth É, Szabó D, et al. Can a brief antepartum preventive group intervention help reduce postpartum depressive symptomatology? Psychotherapy and Psychosomatics. 2012;81:98-107.

Kroenke K, Spitzer RL, Williams JBW. The Patient Health Questionnaire-2: validity of a two-item depression screener. Medical Care. 2003;41:1284-92.

Kropp F, Winhusen T, Lewis D, Hague D, Somoza E. Increasing prenatal care and healthy behaviors in pregnant substance users. Journal of Psychoactive Drugs. 2010;42:73-81.

Kulaga S, Sheehy O, Zargarzadeh AH, Moussally K, Berard A. Antiepileptic drug use during pregnancy: Perinatal outcomes. Seizure. 2011;20:667-72.

Kulin NA, Pastuszak A, Sage SR, Schick-Boschetto B, Spivey G, Feldkamp M, et al. Pregnancy outcome following maternal use of the new selective serotonin reuptake inhibitors: a prospective controlled multicenter study. Journal of the American Medical Association. 1998;279:609-10.

Kumar R. Postpartum psychosis. Bailliere's Clinical Obstetrics and Gynaecology. 1989;3:823-38.

Kumar R, Robson KM, Smith AMR. Development of a self-administered questionnaire to measure material adjustment and material attitudes during pregnancy and after delivery. Journal of Psychosomatic Research. 1984;28:43-51.

Laegreid L, Olegard R, Conradi N, Hagberg G, Wahlstrom J, Abrahamsson L. Congenital malformations and maternal consumption of benzodiazepines: a case-control study. Developmental Medicine & Child Neurology. 1990;32:432-41.

Laegreid L, Hagberg G, Lundberg A. Neurodevelopment in late infancy after prenatal exposure to benzodiazepines - a prospective study. Neuropediatrics. 1992;23:60-7.

Laine K, Heikkinen T, Ekblad U, Kero P. Effects of exposure to selective serotonin reuptake inhibitors during pregnancy on serotonergic symptoms in newborns and cord blood monoamine and prolactin concentrations. Archives of General Psychiatry. 2003;60:720-6.

Lancaster CA, Gold KJ, Flynn HA, Yoo H, Marcus SM, Davis MM. Risk factors for depressive symptoms during pregnancy: a systematic review. American Journal of Obstetrics and Gynecology. 2010;202:5-14.

Lau MA, Segal ZV, Williams JMG. Teasdale's differential activation hypothesis: implications for mechanisms of depressive relapse and suicidal behaviour. Behaviour Research and Therapy. 2004;42:1001-17.

Lau Y, Wang Y, Yin L, Chan KS, Guo X. Validation of the Mainland Chinese version of the Edinburgh Postnatal Depression Scale in Chengdu mothers. International Journal of Nursing Studies. 2010;47:1139-51.

Laurent HK, Leve LD, Neiderhiser J M, Natsuaki MB, Shaw DS, Fisher PA, et al. Effects of prenatal and postnatal parent depressive symptoms on adopted child HPA regulation: independent and moderated influences. Developmental Psychology. 2013;49:876-86.

Lavender T, Walkinshaw SA. Can midwives reduce postpartum psychological morbidity? A randomized trial. Birth. 1998;25:215-19.

Lawrie TA, Hofmeyr GJ, De Jager M, Berk M. Validation of the Edinburgh Postnatal Depression Scale in a cohort of South African women. South African Medical Journal. 1998a;86:1340-4

Lawrie TA, Hofmeyr GJ, De Jager M, Berk M, Paiker J, Viljoen E. A double-blind randomised placebo controlled trial of postnatal norethisterone enanthate: the effect of postnatal depression and serum hormones. British Journal of Obstetrics and Gynaaecology. 1998b;105:1082-90.

Le H-N, Perry DF, Stuart EA. Randomized controlled trial of a preventive intervention for perinatal depression in high-risk latinas. Journal of Consulting and Clinical Psychology. 2011;79:135-41.

Lee DT, Yip SK, Chiu HF, Leung TY, Chan KP, Chau IO, et al. Detecting postnatal depression in Chinese women. Validation of the Chinese version of the Edinburgh Postnatal Depression Scale.British Journal of Psychiatry. 1998;172:433-7.

Leonardou A, ZervasaYM, Papageorgioua CC, Marksb MN, TsartsaraaEC, Antsaklisc A,et al. Validation of the Edinburgh Postnatal Depression Scale and prevalence of postnatal depression at two months postpartum in a sample of Greek mothers. Journal of Reproductive and Infant Psychology. 2009;27:28-39.

Leppee M, Culig J, Eric M, Sijanovic S. The effects of benzodiazepines in pregnancy. Acta Neurologica Belgica. 2010;110:163-67.

Letourneau N, Stewart M, Dennis C-L, Hegadoren K, Duffett-Leger L, Watson B. Effect of home-based peer support on maternal-infant interactions among women with postpartum depression: a randomized, controlled trial. International Journal of Mental Health Nursing. 2011;20:345-57.

Leung SS, Lam TH. Group antenatal intervention to reduce perinatal stress and depressive symptoms related to intergenerational conflicts: a randomized controlled trial. International Journal of Nursing Studies. 2012;49:1391-1402.

Leverton TJ, Elliott SA. Is the EPDS a magic wand? A comparison of the Edinburgh Postnatal Depression Scale and health visitor report as predictors of diagnosis on the present state examination. Journal of Reproductive and Infant Psychology. 2000;18:279-96.

Levine RJ, Hauth JC, Curet LB, Sibai BM, Catalano PM, Morris CD et al. Trial of calcium to prevent preeclampsia. New England Journal of Medicine. 1997;337:69-76.

Levinson-Castiel R, Merlob P, Linder N, Sirota L, Klinger G. Neonatal abstinence syndrome after in utero exposure to selective serotonin reuptake inhibitors in term infants. Archives of Pediatrics and Adolescent Medicine. 2006;160:173-6.

Lewis G, Drife J. Why Mothers Die 2000–2002: The Sixth Report of Confidential Enquiries into Maternal Deaths in the United Kingdom. London: CEMACH/Royal College of Obstetricians and Gynaecologists; 2004.

Lin H-C, Chen IJ, Chen Y-H, Lee H-C, Wu F-J. Maternal schizophrenia and pregnancy outcome: does the use of antipsychotics make a difference? Schizophrenia Research. 2010;116:55-60.

Lin HL, Chen YH, Lin HC. No increase in adverse pregnancy outcomes for women receiving antiepileptic drugs. Journal of Neurology. 2009;256:1742-49.

Lindahl V, Pearson JL, Colpe L. Prevalence of suicidality during pregnancy and the postpartum. Archives of Women's Mental Health. 2005;8:77-87.

Llorente AM, Jensen CL, Voigt RG, Fraley MPH, Berretta LMS, Heird WC. Effect of maternal docosahexaenoic acid supplementation on postpartum depression and information processing. American Journal of Obstetrics and Gynecology. 2003;188:1348-53.

Logsdon MC, Myers JA. Comparative performance of two depression screening instruments in adolescent mothers. Journal of Women's Health. 2010;19:1123-1128.

Lovejoy MC, Graczyk PA, O'Hare E, Neuman G. Maternal depression and parenting behavior: a meta-analytic review. Clinical Psychology Review. 2000;20:561-92.

Lovestone S, Kumar R. Postnatal psychiatric illness: the impact on partners. The British Journal of Psychiatry. 1993;163:210-16.

Lui S, Terplan M, Smith EJ. Psychosocial interventions for women enrolled in alcohol treatment during pregnancy. Cochrane Database of Systematic Reviews. 2008;3:Art. No.: CD006753. DOI: 10.1002/14651858.CD006753.pub2

Lydsdottir LB, Howard LM, Olafsdottir H, Thome M, Tyrfingsson P, Sigurdsson JF. The mental health characteristics of pregnant women with depressive symptoms identified by the Edinburgh Postnatal Depression Scale. Journal of Clinical Psychiatry. 2014. [Epub ahead of print]

Mahmud WM, Awang A, Mohamed MN. Revalidation of the Malay Version of the Edinburgh Postnatal Depression Scale (EPDS) Among Malay Postpartum Women

Attending the Bakar Bata Health Center in AlorSetar, Kedah, North West Of Peninsular Malaysia. The Malaysian Journal of Medical Sciences. 2003;10:71-5.

Makrides M, Gibson RA, McPhee AJ, Yelland L, Quinlivan J, Ryan P.Effect of DHA supplementation during pregnancy on maternal depression and neurodevelopment of young children: a randomized controlled trial. JAMA: The Journal of the American Medical Association. 2010;304:1675-1683.

Malm H, Artama M, Gissler M, Ritvanen A. Selective serotonin reuptake inhibitors and risk for major congenital anomalies. Obstetrics and Gynecology. 2011;118:111-20.

Manber R, Schnyer RN, Allen JJB, Rush JA, Blasey CM. Acupuncture: a promising treatment for depression during pregnancy. Journal of Affective Disorders. 2004;83:89-95.

Manber R, Schnyer RN, Lyell D, Chambers AS, Caughey AB, Druzin M et al. Acupuncture for depression during pregnancy: a randomized controlled trial. Obstetrics and Gynecology. 2010;115:511-20.

Mann R, Gilbody SM. Should methodological filters for diagnostic test accuracy studies be used in systematic reviews of psychometric instruments? A case study involving screening for postnatal depression.Systematic Reviews. 2012;1:9.

Mann T. Clinical Guidelines: Using Clinical Guidelines to Improve Patient Care Within the NHS. London: NHS Executive. 1996.

Mann T. A National Service Framework for Mental Health. Modern Standards and Service Models. London: Department of Health; 1999.

Mann R, Adamson J, Gilbody SM. Diagnostic accuracy of case-finding questions to identify perinatal depression. Canadian Medical Association Journal. 2012:184;e424-430.

Mapp T, Hudson K. Feelings and fears during obstetric emergencies, part 1. British Journal of Midwifery. 2005a;13:30-5.

Mapp T. Feelings and fears post obstetric emergencies, part 2. British Journal of Midwifery. 2005b;13:36-40.

Marais S, Jordaan E, Viljoen D, Olivier L, de Waal J, Poole C. The effect of brief interventions on the drinking behaviour of pregnant women in a high-risk rural South African community: a cluster randomised trial. Early Child Development and Care. 2011;181:463-474.

Maschi S, Clavenna A, Campi R, Schiavetti B, Bernat M, Bonati M. Neonatal outcome following pregnancy exposure to antidepressants: a prospective controlled cohort study. BJOG: An International Journal of Obstetrics and Gynaecology. 2008;115:283-89.

Matthey S. Using the Edinburgh Postnatal Depression Scale to screen for anxiety disorders. Depression and Anxiety. 2008;25:926-31.

Mayet S, Groshkova T, Morgan L, MacCormack T, Strang J. Drugs, alcohol and pregnant women: changing characteristics of women engaging with a specialist perinatal outreach addictions service. Drug Alcohol Review. 2008;27:490-6.

Mazhari S, Nakhaee N. Validation of the Edinburgh Postnatal Depression Scale in an Iranian sample. Archives of Women's Mental Health. 2007;10:293-7.

McAndrew F, Thompson J, Fellows L, Large A, Speed M, Renfrew M. Infant Feeding Survey 2010: Summary. University of Dundee, IFF Research and NHS Information Centre for Health and Social Care. London, NHS Information Centre for Health and Social Care. 2010. Available at: http://doc.ukdataservice.ac.uk/doc/7281/mrdoc/pdf/7281_ifs-uk-2010_report.pdf (accessed 2 July 2014)

McColl H, Dhillon M, Howard LM. A systematic review of the nutritional status of pregnant women with severe mental illness. Archives of Women's Mental Health. 2013;16:39-46

McCreight BS. Perinatal loss: a qualitative study in Northern Ireland. Omega. 2008;57:1-19.

McGrath L, Peters S, Wieck A, Wittkowski A. The process of recovery in women who experienced psychosis following childbirth. BMC Psychiatry. 2013;13:341.

McIntosh E, Barlow J, Davis H, Stewart-Brown S. Economic evaluation of an intensive home visiting programme for vulnerable families: a cost-effectiveness analysis of a public health intervention. Journal of Public Health: Oxford Journal. 2009;31:423-33.

McKenna K, Koren G, Tetelbaum M, Wilton L, Shakir S, Diav-Citrin O, et al. Pregnancy outcome of women using atypical antipsychotic drugs: a prospective comparative study. Journal of Clinical Psychiatry. 2005;66:444-9.

McKnight RF, Adida M, Budge K, Stockton S, Goodwin GM, Geddes JR. Lithium toxicity profile: a systematic review and meta-analysis. Lancet. 2012;379:721-28.

McManus S, Meltzer H, Brugha T, et al. Adult Psychiatric Morbidity in England, 2007. Results of a Household Survey. Health and Social Information Centre, Social Care Statistics. 2009.

Meijssen D, Wolf M-J, Koldewijn K, Houtzager BA, van Wassenaer A, Tronick E, et al. The effect of the infant behavioral assessment and intervention program on mother-infant interaction after very preterm birth. Journal of Child Psychology and Psychiatry. 2010a;51:1287-95.

Meijssen D, Wolf MJ, Koldewijn K, van Wassenaer AG, Kok JH, van Baar AL. Parenting stress in mothers after very preterm birth and the effect of the infant behavioural assessment and intervention program. Child: Care, Health and Development. 2010b;37:195-202.

Meijssen D, Wolf M-J, Koldewijn K, van Baar A, Kok J. Maternal psychological distress in the first two years after very preterm birth and early intervention. Early Child Development and Care. 2011;181:1-11.

Melnyk BM, Feinstein NF, Alpert-Gillis L, Fairbanks E, Crean HF, Sinkin RA, et al. Reducing premature infants' length of stay and improving parents' mental health outcomes with the Creating Opportunities for Parent Empowerment (COPE) neonatal intensive care unit program: a randomized, controlled trial. Pediatrics. 2006;118:e1414-e27.

Meltzer H, Gill B, Petticrew M, Hinds K. OPCS Surveys of Psychiatric Morbidity in Great Britain, Report 1: The Prevalence of Psychiatric Morbidity Among Adults Living in Private Households HMSO: London; 1995a.

Meltzer H, Gill B, Petticrew M, Hinds K. OPCS Surveys of Psychiatric Morbidity in Great Britain, Report 3: Economic Activity and Social Functioning of Adults with Psychiatric Disorders, HMSO: London; 1995b.

Mendelson T, Leis JA, Perry DF, Stuart EA, Tandon SD. Impact of a preventative intervention for perinatal depression on mood regulation, social support, and coping. Archives of Womens Mental Health. 2013;16:211-18.

Meyer EC, Coll CTG, Lester BM, Boukydis CFZ, McDonough SM, et al. Family-based intervention improves maternal psychological well-being and feeding interaction of preterm infants. Pediatrics. 1994;93:241-46.

Micali N, Simonoff E, Treasure J. Pregnancy and post-partum depression and anxiety in a longitudinal general population cohort: the effect of eating disorders and past depression. Journal of Affective Disorders. 2011;131:150-7.

Milgrom J, Ericksen J, Negri L, Gemmill AW. Screening for postnatal depression in routine primary care: properties of the Edinburgh Postnatal Depression Scale in an Australian sample. The Australian and New Zealand Journal of Psychiatry. 2005a;39:833-9.

Milgrom J, Meitz A. Synchrony Coding Scale in Maternal Perceptions of Infant Temperament, Maternal Compentency, Mothering Style and Later Mother–infant Synchrony. Melbourne College of Advanced Education: Australia; 1988.

Milgrom J, Negri LM, Gemmill AW, McNeil M, Martin PR. A randomized controlled trial of psychological interventions for postnatal depression. British Journal of Clinical Psychology. 2005b;44:529-42.

Milgrom J, Schembri C, Ericksen J, Ross J, Gemmill AW. Towards parenthood: an antenatal intervention to reduce depression, anxiety and parenting difficulties. Journal of affective disorders. 2011a;130:385-94.

Milgrom J, Holt CJ, Gemmill AW, Ericksen J, Leigh B, Buist A, et al. Treating postnatal depressive symptoms in primary care: a randomised controlled trial of GP

management, with and without adjunctive counselling. BMC Psychiatry. 2011b;11:95.

Miller LJ. Use of electroconvulsive therapy during pregnancy. Obstetrical and Gynaecological Survey. 1995;50:10-11.

Minkovitz CS, Strobino D, Scharfstein D, et al. Maternal depressive symptoms and children's receipt of health care in the first 3 years of life. Pediatrics. 2005;115:306-14.

Minozzi S, Amato L, Vecchi S, Davoli M. Maintenance agonist treatments for opiate dependent pregnant women. Cochrane Database of Systematic Reviews. 2008;2:Art. No.: CD006318. DOI: 10.1002/14651858.CD006318.pub2

Minozzi S, Amato L, Bellisario C, Ferri M, Davoli M. Maintenance agonist treatments for opiate-dependent pregnant women. Cochrane Database of Systematic Reviews. 2013;12:Art. No.: CD006318. DOI: 10.1002/14651858.CD006318.pub3

Misri S, Kostaras X, Fox D, Kostaras D. The impact of partner support in the treatment of postpartum depression. Canadian Journal of Psychiatry. 2000;45:554-58.

Mitchell AJ, Vaze A, Rao S. Clinical diagnosis of depression in primary care: a meta-analysis. Lancet. 2009;374:609-19.

Mitchell EA, Thompson JM, Stewart AW, et al. Postnatal depression and SIDS: a prospective study. Journal of Paediatrics and Child Health. 1992;28:S13-6.

Mitchell JT. When disaster strikes... The critical incident stress debriefing process. Journal of Emergency Medical Services. 1983;8:36-9.

Mokhber N, Namjoo M, Tara F, Boskabadi H, Rayman MP, Ghayour-Mobarhan M. Effect of supplementation with selenium on postpartum depression: a randomized double-blind placebo-controlled trial. Journal of Maternal-Fetal and Neonatal Medicine. 2011;24:104-108.

Molgaard-Nielsen D, Hviid A. Newer-generation antiepileptic drugs and the risk of major birth defects. Obstetrical and Gynecological Survey. 2011;66:543-4.

Molyneaux E, Poston L, Howard LM. Obesity and mental disorders during pregnancy and postpartum: a systematic review and meta-analysis. Obstetrics and Gynecology. 2014;123:857-67.

Morgan JF, Lacey JH, Sedgwick PM. Impact of pregnancy on bulimia nervosa. The British Journal of Psychiatry. 1999;174:135-40.

Morrell CJ, Spiby H, Stewart P, Walters S, Morgan A. Costs and effectiveness of community postnatal support workers: randomised controlled trial. BMJ. 2000;321:593-8.

Morrell CJ, Warner R, Slade P, Dixon S, Walters S, Paley G, et al. Psychological interventions for postnatal depression: Cluster randomised trial and economic evaluation. The PoNDER trial. Health Technology Assessment. 2009a;13:i-153.

Morrell CJ, Slade P, Warner R, Paley G, Dixon S, Walters SJ, et al. Clinical effectiveness of health visitor training in psychologically informed approaches for depression in postnatal women: pragmatic cluster randomised trial in primary care. BMJ. 2009b;338:a3045.

Morrell CJ, Ricketts T, Tudor K, Williams C, Curran J, Barkham M. Training health visitors in cognitive behavioural and person-centred approaches for depression in postnatal women as part of a cluster randomised trial and economic evaluation in primary care: the PoNDER trial. Primary Health Care Research and Development. 2011;12:11-20.

Morrow J, Russell A, Guthrie E, Parsons L, Robertson I, Waddell R, et al. Malformation risks of antiepileptic drugs in pregnancy: a prospective study from the UK Epilepsy and Pregnancy Register. Journal of Neurology, Neurosurgery and Psychiatry. 2006;77:193-8.

Mozurkewich EL, Clinton CM, Chilimigras JL, Hamilton S, Allbaugh L, Berman D, et al. The Mothers, Omega-3, and Mental Health Study: a double-blind, randomized controlled trial. American Journal of Obstetrics and Gynecology. 2013;208:e1-9.

Mulcahy R, Reay RE, Wilkinson RB, Owen C. A randomised control trial for the effectiveness of group interpersonal psychotherapy for postnatal depression. Archives of Women's Mental Health. 2010;13:125-39.

Mullins SM, Suarez M, Ondersman SJ, Page MC. The impact of motivational interviewing on substance abuse treatment retention: A randomised control trial of women involved with child welfare. Journal of Substance Abuse Treatment. 2004;27:51-8.

Munk-Olsen T, Laursen TM, Pedersen CB, Mors O, Mortensen PB. New parents and mental disorders: population-based register study. Journal of the American Medical Association. 2006;296:2582-9.

Munk-Olsen T, Laursen TM, Pedersen CB, Mors O, Mortensen PB. Family and partner psychopathology and the risk of postpartum mental disorders. Journal of Clinical Psychiatry. 2007;68:1947-53.

Munk-Olsen T, Laursen TM, Mendelson T, Pedersen CB, Mors O, Mortensen PB. Risks and predictors of readmission for a mental disorder during the postpartum period. Archive of General Psychiatry. 2009;66:189-95.

Muñoz RF, Le H-N, Ippen CG, Diaz MA, Urizar Jr. GG, Soto J, et al. Prevention of postpartum depression in low-income women: development of the Mamas y Bebes/Mothers and Babies course. Cognitive and Behavioral Practice. 2007;14:70-83.

Murray D, Cox JL. Screening for depression during pregnancy with the Edinburgh Depression Scale (EDDS).Journal of Reproductive and Infant Psychology. 1990;8:99-107.

Murray L, Carothers AD.The validation of the Edinburgh Postnatal Depression Scale on a community sample. British Journal of Psychiatry. 1990a;157:288-90.

Murray L, Fiori-Cowley A, Hooper R, Cooper P.The impact of postnatal depression and associated adversity on early mother-infant interactions and later infant outcome. Child Development. 1996;67:2512-26.

Murray L, Cooper PJ, Wilson A, Romaniuk H. 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. 2003;182:420-7.

Mustafa RA, Santesso N, Brozek J, Akl EA, Walter SD, Norman G et al. The GRADE approach is reproducible in assessing the quality of evidence of quantitative evidence syntheses. Journal of Clinical Epidemiology. 2013;66:736-42.

Muzik M, Klier CM, Rosenblum KL, Holzinger A, Umek W, Katschnig H. Are commonly used self-report inventories suitable for screening postpartum depression and anxiety disorders? Acta Psychiatrica Scandinavica. 2000;102:71-3.

Myles N, Newall H, Ward H. Large M. Systematic meta-analysis of individual selective serotonin reuptake inhibitor medications and congenital malformations. Australian and New Zealand Journal of Psychiatry. 2013:47;1002-12.

Nager A, Johansson LM, Sundquist K. Are sociodemographic factors and year of delivery associated with hospital admission for postpartum psychosis? A study of 500,000 first-time mothers. Acta Psychiatrica Scandinavica. 2005;112:47-53.

Nair P, SchulerME, Kettinger L, Harrington D. Cumulative environmental risk in substance abusing women: parenting stress, child abuse potential, and development. Pediatric Research. 2002;51:186A.

Nair P, Schuler ME, Black MM, Kettinger L, Harrington D. Cumulative environmental risk in substance abusing women: early intervention, parenting stress, child abuse potential and child development. Child Abuse and Neglect. 2003;29:997-1017.

Nair P, Black MM, Ackerman JP, Schuler ME, Keane VA. Children's cognitive-behavioral functioning at age 6 and 7: prenatal drug exposure and caregiving environment. Ambulatory Pediatrics. 2008;8:154-62.

National Clinical Guidline Centre. Patient Experience in Adult NHS Services: Improving the Experience of Care for People Using Adult NHS Services. NICE clinical guideline 138. London: NICE; 2012. Available from: www.nice.org.uk/guidance/cg138/

NCCMH. Post-Traumatic Stress Disorder: The Management of PTSD in Adults and Children in Primary and Secondary Care. Clincal guideline 26. Leicester and London: The British Psycholgical Society and The Royal College of Psychiatrists; 2005. [Full guideline]

NCCMH. Bipolar Disorder: The Management of Bipolar Disorder in Adults, Children and Adolescents, in Primary and Secondary Care. Clincal guideline 38. Leicester and London: The British Psycholgical Society and The Royal College of Psychiatrists; 2006. [Full guideline]

NCCMH. Antenatal and Postnatal Mental Health: Clinical Management and Service Guidance. Clinical guideline 45. Leicester and London: The British Psycholgical Society and The Royal College of Psychiatrists; 2007. [Full guideline]

NCCMH. Depression (Update): The Treatment and Management of Depression in Adults. Clinical guideline 90. Leicester and London: The British Psycholgical Society and The Royal College of Psychiatrists; 2009. [Full guideline]

NCCMH. Common Mental Health Disorders: The NICE Guideline on Identification and Pathways to Care. Clinical guideline 123. Leicester & London: The British Psychological Society and the Royal College of Psychiatrists; 2011a. [Full guideline]

NCCMH. Generalised Anxiety Disorder in Adults: Management in Primary, Secondary and Community Care. Clinical guideline 113. Leicester & London: The British Psychological Society and the Royal College of Psychiatrists; 2011. [Full guideline]

NCCMH. Service User Experience in Adult Mental Health: Improving the experience of Care for People Using Adult NHS Mental Health Services. Clinical guideline 136. Leicester and London: The British Psychological Society and the Royal College of Psychiatrists; 2012. [Full guideline]

NCCMH. Psychosis and Schizophrenia in Adults (Update): The NICE Guideline on Treatment and Management. Clinical guideline 176. Leicester and London: The British Psycholgical Society and The Royal College of Psychiatrists; 2014. [Full guideline]

Neugebauer R, Kline J, Markowitz JC, Bleiberg KL, Baxi L, Rosing MA, et al. Pilot randomized controlled trial of interpersonal counseling for subsyndromal depression following miscarriage. Journal of Clinical Psychiatry. 2006;67:1299-304.

Newnham CA, Milgrom J, Skouteris H. Effectiveness of a modified mother-infant transaction program on outcomes for preterm infants from 3 to 24 months of age. Infant Behavior and Development. 2009;32:17-26.

Newham JJ, Thomas SH, MacRitchie K, McElhatton PR, McAllister-Williams RH. Birth weight of infants after maternal exposure to typical and atypical antipsychotics: prospective comparison study. British Journal of Psychiatry. 2008;192:333-7.

NHS Business Services Authority & NHS Prescription Services (2014) Electronic Drug Tariff for England and Wales, June 2014. Compiled on behalf of the Department of Health. Available from: http://www.ppa.org.uk/ppa/edt_intro.htm [accessed 4 June 2014].NICE. Guidance on the use of electroconvulsive therapy. NICE technology appraisal 59. London: NICE; 2003.

NICE. Eating Disorders: Core Interventions in the Treatment and Management of Anorexia Nervosa, Bulimia Nervosa and Related Eating Disorders. NICE clinical guideline 9. London: NICE; 2004a. Available from: www.nice.org.uk/guidance/CG9

NICE. Post-Traumatic Stress Disorder: The Management of PTSD in Adults and Children in Primary and Secondary Care. NICE clinical guideline 26. London: NICE; 2005a. Available from: www.nice.org.uk/guidance/CG26

NICE. Obsessive-compulsive Disorder: Core Interventions in the Treatment of Obsessive-compulsive Disorder and Body Dysmorphic Disorder. NICE clinical guideline 31. London: NICE; 2005b. Available from: www.nice.org.uk/guidance/CG31

NICE. Bipolar Disorder: The Management of Bipolar Disorder in Adults, Children and Adolescents, in Primary and Secondary Care. NICE clinical guideline 38. London: NICE; 2006. Available from: www.nice.org.uk/guidance/CG38

NICE. Antenatal and Postnatal Mental Health: Clinical Management and Service Guidance. NICE clinical guidance 45. London: NICE; 2007a. Available from: http://www.nice.org.uk/guidance/cg45

NICE. Drug Misuse: Opioid Detoxification. NICE clinical guideline 52. London: NICE; 2007b. Available from: www.nice.org.uk/guidance/CG52

NICE. Social Value Judgements. Principles for the Development of NICE Guidance. London: NICE; 2008. Available from: www.nice.org.uk/media/C18/30/SVJ2PUBLICATION2008.pdf

NICE. Depression in Adults: The Treatment and Management of Depression in Adults. NICE clinical guideline 90. London: NICE; 2009a. Available from: www.nice.org.uk/guidance/CG90

NICE. Depression in Adults with a Chronic Physical Health Problem: Treatment and Management. NICE clinical guideline 91. London: NICE; 2009b. Available from: www.nice.org.uk/guidance/CG91

NICE. Pregnancy and Complex Social Factors: A Model for Service Provision for Pregnant Women with Complex Social Factors. NICE clinical guideline 110. London: NICE; 2010. Available from: www.nice.org.uk/guidance/cg110

NICE. Generalised Anxiety Disorder in Adults: Management in Primary, Secondary and Community Care. NICE clinical guideline 113. London: NICE; 2011a. Available from: www.nice.org.uk/guidance/CG113

NICE. Common Mental Health Disorders: Identification and Pathways to Care. NICE clinical guideline 123. London: NICE; 2011b. Available from: www.nice.org.uk/guidance/CG123

NICE. Alcohol-Use Disorders: Diagnosis, Assessment and Management of Harmful Drinking and Alcohol Dependence. NICE clinical Guideline 115. London: NICE; 2011c. Available from: www.nice.org.uk/guidance/CG115

NICE. Service User Experience in Adult Mental Health: Improving the Experience of Care for People Using Adult NHS Mental Health Services. NICE Clinical Guideline 136. London: NICE; 2011d. Available from: guidance.nice.org.uk/guidance/CG136

NICE. Caesarean Section (Update). Costing report. Implementing NICE guidance. NICE clinical guideline 132. London: NICE; 2011e. Available from: www.nice.org.uk/nicemedia/live/13620/57173/57173.pdf. 2011

NICE. The Guidelines Manual. London: NICE; 2012a. Available from: publications.nice.org.uk/the-guidelines-manual-pmg6

NICE. Patient Experience in Adult NHS Services: Improving the Experience of Care for People Using Adult NHS Services. London: NICE; 2012b. Available from: www.nice.org.uk/guidance/cg138

NICE. Social Anxiety Disorder: Recognition, Assessment and Treatment. NICE clinical guideline 159. London: NICE; 2013a. Available from: www.nice.org.uk/guidance/CG159

NICE. Guide to the Methods of Technology Appraisal 2013. London: NICE; 2013b. Available from: http://publications.nice.org.uk/pmg9

NICE. Psychosis and Schizophrenia in Adults: Treatment and Management. NICE clinical guideline 178. London: NICE; 2014a. Available from: www.nice.org.uk/guidance/CG178

NICE. Bipolar Disorder: Management of Bipolar Disorder in Adults, Children and Adolescents, in Primary and Secondary Care. NICE clinical guideline 38. London: NICE; 2014b. Available from: www.nice.org.uk/CG185

Nicholls K, Ayers S. Childbirth-related post-traumatic stress disorder in couples: a qualitative study. British Journal of Health Psychology. 2007;12:491-509.

Nikčević AV, Kuczmierczyk AR, Nicolaides KH. The influence of medical and psychological interventions on women's distress after miscarriage. Journal of Psychosomatic Research. 2007;63:283-90.

Nonacs R, Cohen LS. Postpartum mood disorders: diagnosis and treatment guidelines. The Journal of Clinical Psychiatry. 1998;59:34-40

Norman E, Sherburn M, Osborne RH, Galea MP. An exercise and education program improves well-being of new mothers: a randomized controlled trial. Physical therapy. 2010;90:348-55.

Nulman I, Rovet J, Stewart DE, Wolpin J, Gardner HA, Theis JGW, et al. Neurodevelopment of children exposed in utero to antidepressant drugs. New England Journal of Medicine. 1997;336:258-62.

Nulman I, Rovet J, Stewart DE, Wolpin J, Pace-Asciak P, Shuhaiber S, et al. Child development following exposure to tricyclic antidepressants or fluoxetine

throughout fetal life: a prospective, controlled study. The American Journal of Psychiatry. 2002;159:1889-95.

Nylen KJ, Williamson JA, O'Hara MW, Watson D, Engeldinger J. Validity of somatic symptoms as indicators of depression in pregnancy. Archive of Women's Mental Health. 2013;16:203-10.

O'Brien L, Laporte A, Koren G. Estimating the economic costs of antidepressant discontinuation during pregnancy. Canadian Journal of Psychiatry. 2009;54:399-408.

O'Connor M, Whaley S. Brief intervention for alcohol use by pregnant women. American Journal of Public Health. 2007;97:252-8.

O'Connor TG, Heron J, Golding J, Beveridge M, Glover V. Maternal antenatal anxiety and children's behavioural/emotional problems at 4 years. Report from the Avon Longitudinal Study of Parents and Children. British Journal of Psychiatry. 2002;180:502-8.

O'Connor TG, Heron J, Golding J, Glover V. Maternal antenatal anxiety and behavioural/emotional problems in children: a test of a programming hypothesis. Journal of child psychology and psychiatry, and allied disciplines. 2003a;44:1025-36.

O'Connor TG, Heron J, Golding J, Glover V, ALSPAC Study team. Maternal antenatal anxiety and behavioural/emotional problems in children: a test of a programming hypothesis. Journal of Child Psychology and Psychiatry, and Allied Disciplines. 2003b;44:1025-36.

O'Hara MW, Swain AM. Rates and risk of postpartum depression – a meta-analysis. International Review of Psychiatry. 1996;8:37-54.

O'Hara MW, Stuart S, Gorman LL, Wenzel A. Efficacy of interpersonal psychotherapy for postpartum depression. Archives of General Psychiatry. 2000;57:1039-45.

O'Higgins M, St. James Roberts I, Glover V. Postnatal depression and mother and infant outcomes after infant massage. Journal of Affective Disorders. 2008;109:189-92.

O'Leary CM, Nassar N, Kurinczuk JJ, Bower C. The effect of maternal alcohol consumption on fetal growth and preterm birth. BJOG: An International Journal of Obstetrics and Gynaecology. 2009;116:390-400.

O'Mahen HA, Woodford J, McGinley J, Warren FC, Richards DA, Lynch TR, et al., Internet-based behavioral activation - treatment for postnatal depression (Netmums): a randomized controlled trial. Journal of Affective Disorders. 2013a;150:814-22.

O'Mahen H, Himle JA, Fedock MA, Henshaw E, Flynn H. A pilot randomised controlled trial of cognitive behavioural therapy for perinatal depression adapted for women with low incomes. Depression and Anxiety. 2013b;30:679-87.

O'Mahen HA, Richards DA, Woodford J, Wilkinson E, McGinley J, Taylor RS, et al. Netmums: a phase II randomized controlled trial of a guided internet behavioural activation treatment for postpartum depression. Psychological Medicine. 2013c;23:1-15. [Epub ahead of print]

O'Mahen HA, Woodford J, McGinley J, Warren FC, Richards DA, Lynch TR, et al., Internet-based behavioral activation - treatment for postnatal depression (Netmums): a randomized controlled trial. Journal of Affective Disorders. 2013;150:814-22.

O'Neill K, Baker A, Cooke M, Collins E, Heather N, Wodak A. Evaluation of a cognitive-behavioural intervention for pregnant injecting drug users at risk of HIVinfection. Addiction. 1996;91:1115-25.

Oates, M. (2000) Perinatal Maternal Mental Health Services. Council Report CR88.

Oates M, Cantwell R. Chapter 11: Deaths from psychiatric causes in Saving Mothers' Lives. Reviewing maternal deaths to make motherhood safer: 2006-2008 in The Eighth Report of the Confidential Enquiries into Maternal Deaths in the United Kingdom (CMACE). British Journal of Obstetrics and Gynaecology. 2011;118:1-203

Oberlander TF, Warburton W, Misri S, Aghajanian J, Hertzman C. Neonatal outcomes after prenatal exposure to selective serotonin reuptake inhibitor antidepressants and maternal depression using population-based linked health data. Archives of General Psychiatry. 2006;63:898-906.

Oberlander TF, Warburton W, Misri S, Riggs W, Aghajanian J, Hertzman C. Major congenital malformations following prenatal exposure to serotonin reuptake inhibitors and benzodiazepines using population-based health data. Birth Defects Research Part B - Developmental and Reproductive Toxicology. 2008;83:68-76.

Office for National Statistics. Birth Statistics: Review of the Registrar General on births and patterns of family building in England and Wales, 2004, Series FM1 33. London: Office for National Statistics; 2006. Available at: http://www.statistics.gov.uk [accessed June 2006]

Office for National Statistics. Statistical Bulletin: Births in England and Wales 2012. London: Office for National Statistics; 2012a. Available from: http://www.ons.gov.uk/ons/rel/vsob1/birth-summary-tables--england-and-wales/2012/stb-births-in-england-and-wales-2012.html [accessed March 2014].

Office for National Statistics. Statistical Bulletin: Infant and Perinatal Mortality in England and Wales by Social and Biological Factors. London: Office for National Statistics; 2012b. Available from: http://www.ons.gov.uk/ons/rel/child-health/infant-and-perinatal-mortality-in-england-and-wales-by-social-and-biological-factors/2011/stb-infant-and-perinatal-mortality--2011.html [accessed March 2014]

Office for National Statistics. Statistical Bulletin: Gestation-specific Infant Mortality in England and Wales , 2011. London: Office for National Statistics; 2013. Available

from: http://www.ons.gov.uk/ons/dcp171778_329137.pdf [accessed November 2014]

Olde E, Van Der Hart O, Kleber R, van Son M. Posttraumatic stress following childbirth: a review. Clinical Psychology Review.2006;26:1-16.

Ondersma SJ, Winhusen T, Lewis DF. Pre-treatment change in a randomized trial with pregnant substance-abusing women in community-based outpatient treatment. Contemporary Clinical Trials. 2012;33:1074-9.

Ondersma SJ, Svikis DS, Thacker LR, Beatty JR, Lockhart N. Computer-delivered screening and brief intervention (e-SBI) for postpartum drug use: a randomized trial. Journal of Substance Abuse Treatment. 2014;46:52-9.

Onozawa K, Glover V, Adams D, Modi N, Kumar RC. Infant massage improves mother-infant interaction for mothers with postnatal depression. Journal of Affective Disorders. 2001;63:201-7.

Oram S, Stockl H, Busza J, Howard LM, Zimmerman C. Systematic review of the prevalence and risk of violence and the physical, mental and sexual health problems associated with human trafficking. PLOS Medicine. 2012;9:5.

Ornoy A, Cohen E. Outcome of children born to epileptic mothers treated with carbamazepine during pregnancy. Archives of disease in childhood. 1996;75:517-20.

Ornoy A, Arnon J, Shechtman S, Moerman L, Lukashova I. Is benzodiazepine use during pregnancy really teratogenic? Reproductive Toxicology. 1998;12:511-15.

Ortiz Collado MA, Saez M, Favrod J, Hatem M. Antenatal psychosomatic programming to reduce postpartum depression risk and improve childbirth outcomes: a randomized controlled trial in Spain and France. BMC Pregnancy and Childbirth. 2014;14:22.

Osterman RL, Dyehouse J. Effects of a motivational interviewing intervention to decrease prenatal alcohol use. Western Journal of Nursing Research. 2012;34:434-54.

Osterman RL, Carle AC, Ammerman RT, Gates D. Single-session motivational intervention to decrease alcohol use during pregnancy. Journal of Substance Abuse Treatment. 2014;47:10-19.

Parvin A, Jones CE, Hull SA. Experiences and understandings of social and emotional distress in the postnatal period among Bangladeshi women living in Tower Hamlets. Family Practice. 2004;21:254-60.

Pastuszak A, Milich V, Chan S, Chu J, Koren G. Prospective assessment of pregnancy outcome following first trimester exposure to benzodiazepines. Canadian Journal of Clinical Pharmacology. 1996;3:167-71.

Patel S, Wittkowski A, Fox JR, Wieck A. An exploration of illness beliefs in mothers with postnatal depression. Midwifery. 2013;29:682-9.

Patrick SW, Schumacher RE, Benneyworth BD, Krans EE, McAllister JM, Davis MM. Neonatal abstinence syndrome and associated health care expenditures: United States, 2000-2009. The Journal of the American Medical Association. 2012;307:1934-40.

Paulden M, Palmer S, Hewitt C, Gilbody S. Screening for postnatal depression in primary care: Cost effectiveness analysis. BMJ. 2009;339:b5203.

Paulson JF, Bazemore SD. Prenatal and postpartum depression in fathers and its association with maternal depression. Journal of the American medical Association. 2010;303:1961-9.

Pawlby S, Hay DF, Sharp D, Waters CS, O'Keane V. Antenatal depression predicts depression in adolescent offspring: prospective longitudinal community-based study. Journal of Affective Disorders. 2009;113:236-43.

Pearson RM, Evans J, Kounali D, Lewis G, Heron J, Ramchandani PG, et al. Maternal depression during pregnancy and the postnatal period: risks and possible mechanisms for offspring depression at age 18 years. JAMA Psychiatry. 2013a;70: 1312-19.

Pearson RM, O'Mahen H, Burns A, Bennert K, Shepherd C, Baxter H, et al. The normalisation of disrupted attentional processing of infant distress in depressed pregnant women following cognitive behavioural therapy. Journal of Affective Disorders. 2013b;145:208-13.

Pedersen LH, Henriksen TB, Vestergaard M, Olsen J, Bech BH. Selective serotonin reuptake inhibitors in pregnancy and congenital malformations: population based cohort study. British Medical Journal. 2009;339:735.

Peng M, Gao K, Ding Y, Ou J, Calabrese JR, Wu R, et al. Effects of prenatal exposure to atypical antipsychotics on postnatal development and growth of infants: a case-controlled, prospective study. Psychopharmacology. 2013;228:577-84.

Perez-Blasco J, Viguer P, Rodrigo MF. Effects of a mindfulness-based intervention on psychological distress, well-being, and maternal self-efficacy in breast-feeding mothers: results of a pilot study. Archives of Womens Mental Health. 2013;16:227-36.

Petrou S, Cooper P, Murray L, Davidson LL. Economic costs of post-natal depression in a high-risk British cohort. The British Journal of Psychiatry. 2002;181:505-12.

Petrou S, Cooper P, Murray L, Davidson LL. Cost-effectiveness of a preventive counseling and support package for postnatal depression. International Journal of Technology Assessment in Health Care. 2006;22:443-53.

Phillips J, Charles M, Sharpe L, Matthey S. Validation of the subscales of the Edinburgh Postnatal Depression Scale in a sample of women with unsettled infants.Journal of Affective Disorders. 2009;118:101-12.

Phipps MG, Raker CA, Ware CF, Zlotnick C. Randomized controlled trial to prevent postpartum depression in adolescent mothers. American Journal of Obstetrics and Gynecology. 2013;208: 192.e1-6.

Pinheiro RT, Botella L, Quevedo LDA, Pinheiro KAT, Jansen K, Osório CM, et al. Maintenance of the effects of cognitive behavioural and relational constructivist psychotherapies in the treatment of women with postpartum depression: a randomized clinical trial. Journal of Constructuvist Psychology. 2014;27:59-68.

Pitanupong J, Liabsuetrakul T, Vittayanont A. Validation of the Thai Edinburgh Postnatal Depression Scale for screening postpartum depression. Psychiatry Research. 2007;149:253-9.

Plant DT, Barker ED, Waters CS, Pawlby S, Pariante CM. Intergenerational transmission of maltreatment and psychopathology: the role of antenatal depression. Psychological Medicine. 2013;43:519-28.

Popova S, Lange S, Burd L, Rehm J. Canadian children and youth in care: he cost of fetal alcohol spectrum disorder. Child and Youth Care Forum. 2014;43:83-96.

Pottegard A, Hallas J, Andersen JT, Lokkegaard ECL, Dideriksen D, Aagaard L. First-trimester Exposure to Methylphenidate: A population-Based Cohort Study. Journal of Clinical Psychiatry. 2014;75:e88-93.

PANDA. Cost of perinatal depression in Australia. Sydney: Deloitte Access Economics; 2012.

Prendergast J, Austin M-P. Early childhood nurse-delivered cognitive behavioural counselling for post-natal depression. Australasian Psychiatry. 2001;9:255-9.

Quinlivan JA, Box H, Evans SF. Postnatal home visits in teenage mothers: a randomised controlled trial. Lancet. 2003;361:893-900.

Rådestad I, Surkan PJ, Steineck G, Cnattingius S, Onelöv E, et al. Long-term outcomes for mothers who have or have not held their stillborn baby. Midwifery. 2009a;25:422-9.

Rådestad I, Säflund K, Wredling R, Onelöv E, Steineck G. Holding a stillborn baby: mothers' feelings of tenderness and grief. British Journal of Midwifery. 2009b;17:178-80.

Rahman A, Malik A, Sikander S, Roberts C, Creed F. Cognitive behaviour therapy-based intervention by community health workers for mothers with depression and their infants in rural Pakistan: a cluster-randomised controlled trial. Lancet. 2008;372:902-9.

Rai D, Lee BK, Dalman C, Golding J, Lewis G, Magnusson C. Parental depression, maternal antidepressant use during pregnancy, and risk of autism spectrum disorders: population based case-control study. British Medical Journal. 2013;346:f2059.

Ramos E, St-Andre M, Rey E, Oraichi D, Berard A. Duration of antidepressant use during pregnancy and risk of major congenital malformations. British Journal of Psychiatry. 2008;192:344-50.

Ramsay R. Postnatal depression. Lancet. 1993;341:1358.

Ravn IH, Smith L, Smeby NA, Kynoe NM, Sandvik L, Bunch EH, et al. Effects of early mother-infant intervention on outcomes in mothers and moderately and late preterm infants at age 1 year: a randomized controlled trial. Infant Behavior and Development. 2012;35:36-47.

Raymond JE. 'Creating a safety net' women's experiences of antenatal depression and their identification of helpful community support and services during pregnancy. Midwifery. 2009;25:39-49.

Rees AM, Austin MP, Parker GB. Omega-3 fatty acids as a treatment for perinatal depression: Randomized double-blind placebo-controlled trial. Australian and New Zealand Journal of Psychiatry. 2008;42:199-205.

Regmi S, Sligl W, Carter D, Grut W, Seear M. A controlled study of postpartum depression among Nepalese women: validation of the Edinburgh Postpartum Depression Scale in Kathmandu. Tropical Medicine and International Health. 2002;7:378-82.

Reis M, Kallen B. Maternal use of antipsychotics in early pregnancy and delivery outcome. Journal of clinical psychopharmacology. 2008;28:279-88.

Reynolds KD, Coombs DW, Lowe JB, Peterson PL, Gayoso E. Evaluation of a self-help program to reduce alcohol consumption among pregnant women. International Journal of the Addictions. 1995;30:427-43.

Rihtman T, Parush S, Ornoy A. Developmental outcomes at preschool age after fetal exposure to valproic acid and lamotrigine: Cognitive, motor, sensory and behavioral function. Reproductive Toxicology. 2013;41:115-25.

Robertson E, Grace S, Wallington T, Stewart DE. Antenatal risk factors for postpartum depression: a synthesis of recent literature. General Hospital Psychiatry. 2004;26:289-95.

Robertson E, Jones I, Haque S, Holder R, Craddock N. Risk of puerperal and non-puerperal recurrence of illness following bipolar affective puerperal (post-partum) psychosis. British Journal of Psychiatry 2005;186:258-9.

Robertson E, Lyons A. Living with puerperal psychosis: a qualitative analysis. Psychology and Psychotherapy. 2003;76:411-31.

Robledo-Colonia AF, Sandoval-Restrepo N, Mosquera-Valderrama YF, Escobar-Hurtado C, Ramirez-Velez R. Aerobic exercise training during pregnancy reduces depressive symptoms in nulliparous women: a randomised trial. Journal of Physiotherapy. 2012;58:9-15.

Rodriguez-Pinilla E, Arroyo I, Fondevilla J, Garcia MJ, Martinez-Frias ML. Prenatal exposure to valproic acid during pregnancy and limb deficiencies: a case-control study. American Journal of Medical Genetics. 2000;90:376-81.

Rodriguez-Pinilla E, Mejias C, Prieto-Merino D, Fernandez P, Martinez-Frias ML. Risk of hypospadias in newborn infants exposed to valproic acid during the first trimester of pregnancy: a case-control study in Spain. Drug Safety. 2008;31:537-43.

Rogers CR. The necessary and sufficient conditions of therapeutic personality change. Journal of Consulting Psychology. 1957;21:95-103.

Roman LA, Gardiner JC, Lindsay JK, Moore JS, Luo Z, Baer LJ, et al. Alleviating perinatal depressive symptoms and stress: A nurse-community health worker randomized trial. Archives of Women's Mental Health. 2009;12:379-91.

Ross LE, McLean LM. Anxiety disorders during pregnancy and the postpartum period: a systematic review. Journal of Clinical Psychiatry. 2006;67:1285-98.

Ross LE, Dennis C-L. The prevalence of postpartum depression among women with substance use, an abuse history, or chronic illness: a systematic review. Journal of Women's Health. 2009;18:475-86.

Ross LE, Grigoriadis S, Mamisashvili L, Vonderporten EH, Roerecke M, Rehm J, et al. Selected pregnancy and delivery outcomes after exposure to antidepressant medication: A systematic review and meta-analysis. Psychiatry. Journal of the American Medical Association. 2013:70; 436-443

Rouhe H, Salmela-Aro K, Gissler M, Halmesmäki E, Saisto T. Mental health problems common in women with fear of childbirth. BJOG: An International Journal of Obstetrics and Gynaecology. 2011;118:1104-11.

Rouhe, H., K. Salmela-Aro, et al. Obstetric outcome after intervention for severe fear of childbirth in nulliparous women - Randomised trial. BJOG: An International Journal of Obstetrics and Gynaecology. 2012;120: 75-84.

Rowel D, Jayawardena P, Fernando N.Validation of the Sinhala translation of Edinburgh Postnatal Depression Scale.The Ceylon Medical Journal. 2008;53:10-3.

Royal College of Physicians. Passive Smoking and Children: A Report by the Tobacco Advisory Group of the Royal College of Physicians. London: Royal College of Physicians; 2010.

Royal College of Physicians, Royal College of Psychiatrists. Smoking and Mental Health: a Joint Report by the Royal College of Physicians and the Royal College of Psychiatrists. London: Royal College of Physicians; 2013.

Royal College of Psychiatrists' Public Education Editorial Board. Postnatal Depression. [Leaflet]. London: Royal College of Psychiatrists; 2014. Available from: http://www.rcpsych.ac.uk/healthadvice/problemsdisorders/postnataldepression.aspx

Rubertsson C, Börjesson K, Berglund A, Josefsson A, Sydsjö G. The Swedish validation of Edinburgh Postnatal Depression Scale (EPDS) during pregnancy. Nordic Journal of Psychiatry. 2011;65:414-18.

Rubovits, P. Project CHILD: An intervention program for psychotic mothers and their young children. In: Gopfert M, Webster J, Seeman MV. Parental Psychiatric Disorder: Distressed Parents and their Families Cambridge: Cambridge University Press; 1996:161-9.

Russell EJ, Fawcett JM, Mazmanian D. Risk of obsessive-compulsive disorder in pregnant and postpartum women: a meta-analysis. Journal of Clinical Psychiatry. 2013;74:377-85.

Ryding EL, Wijma B, Wijma K, Rydhström H. Fear of childbirth during pregnancy may increase the risk of emergency caesarean section. Acta Obstetricia et Gynecologica Scandinavica. 1998;77:542-7.

Ryninks K, Roberts-Collins C, McKenzie-McHarg K, Horsch A. Mothers' experience of their contact with their stillborn infant: an interpretative phenomenological analysis. BMC Pregnancy and Childbirth. 2014;14:203.

Sadowski A, Todorow M, Brojeni PY, Koren G, Nulman I. Pregnancy Outcomes following Maternal exposure to Second-generation antipsychotics given with other psychotropic drugs: a cohort study. BMJ Open. 2013;3. pii: e003062. doi: 10.1136/bmjopen-2013-003062.

Saisto T, Halmesmaki E. Fear of childbirth: a neglected dilemma. Acta Obstetricia et Gynecologica Scandinavica. 2003;82:201-8.

Saisto T, Salmela-Aro K, Nurmi J, Könönen T, Halmesmäki E. A randomized controlled trial of intervention in fear of childbirth. Acta Obstet Gyn Scan. 2001;98: 820-6.

Salmela-Aro K, Read S, Rouhe H, Halmesmäki E, Toivanen RM, Tokola MI, et al. Promoting positive motherhood among nulliparous pregnant women with an intense fear of childbirth: RCT intervention. Journal of Health Psychology. 2012;17:520-34.

Salomonsson B, Sandell R. A randomized controlled trial of mother-infant psychoanalytic treatment: I. outcomes on self-report questionnaires and eternal ratings. Infant Mental Health Journal. 2011;32:207-31.

Samren EB, Van Duijn CM, Christiaens GCML, Hofman A, Lindhout D. Antiepileptic drug regimens and major congenital abnormalities in the offspring. Annals of Neurology. 1999;46:739-46.

Sanderson CA, Cowden B, Hall DMB, Taylor EM, Carpenter RG, Cox JL. Is postnatal depression a risk factor for sudden infant death? The British Journal of General Practice. 2002;52:636-40.

Santos IS, MatijasevichI A, Franck TavaresI B,Barros AJD,PicininiBotelhoI I, Lapoll C, et al. Validation of the Edinburgh Postnatal Depression Scale (EPDS) in a sample of mothers from the 2004 Pelotas Birth Cohort Study. Cadernos de Saude Publica. 2007;23:2577-88.

Sapin C, Fantino B, Nowicki ML, Kind P. Usefulness of EQ-5D in assessing health status in primary care patients with major depressive disorder. Health and Quality of Life Outcomes. 2004;2:20.

Schuler ME, Nair P, Black MM, Kettinger L. Mother-infant interaction: effects of a home intervention and ongoing maternal drug use. Journal of Clinical Child Psychology. 2000;29:424-31.

Schuler ME, Nair P, Black MM. Ongoing maternal drug use, parenting attitudes, and a home intervention: effects on mother-child interaction at 18 months. Journal of Developmental and Behavioral Pediatrics. 2002a;23:87-94.

Schuler ME, Nair P, Keane V. Developmental outcomes of drug-exposed children: effects of a home intervention and ongoing maternal drug use at 3 years. Pediatric Research. 2002b;51:190A.

Schuler ME, Nair P, Kettinger L. Drug-exposed infants and developmental outcome: effects of a home intervention and ongoing maternal drug use. Archives of Pediatrics and Adolescent Medicine. 2003;157:133-8.

Schumacher M, Zubaran C, White G. Bringing birth-related paternal depression to the fore. Women and Birth. 2008;21:65-70.

Schünemann HJ, Best D, Vist G, Oxman A, GRADE Working Group. Letters, numbers, symbols and words: how to communicate grades of evidence and recommendations. Canadian Medical Association Journal. 2003;169:677-80.

Schünemann H, Brożek J, Oxman A, eds. GRADE handbook for grading quality of evidence and strength of recommendation. Version 3.2 [updated March 2009]. The GRADE Working Group, 2009. Available from http://www.cc-ims.net/gradepro.

Scott S, Knapp M, Henderson J, Maughan B. Financial cost of social exclusion: follow up study of antisocial children into adulthood. British Medical Journal. 2001;323:191-4.

Seaton SE, Field DJ, Draper ES, Manketelow BN, Smith GCS, Springett A, et al. Reproductive medicine, obstetrics and gynaecology - research: socioeconomic inequalities in the rate of stillbirths by cause: a population-based study. BMJ Open. 2012;2:3.

Segal Z, Teasdale J, Williams M. Mindfulness-Based Cognitive Therapy for Depression. New York: Guilford Press; 2002.

Sen DM. A randomized controlled trial of midwife-led twin antenatal program – The Newcastle twin study [thesis]. Newcastle-upon-Tyne: University of Newcastle; 2006.

Seng JS, Rauch SA, Resnick H, Reed CD, King A, Low LK. Exploring posttraumatic stress disorder symptom profile among pregnant women. Journal of Psychosomatic Obstetrics and Gynecology. 2010;31:176-87.

Shakespeare J, Blake f,Garcia J. A qualitative study of the acceptability of routine screening of postnatal women using the Edinburgh Postnatal Depression Scale. The British Journal of General Practice. 2003;53:614-19.

Shakespeare J, Blake F, Garcia J. How do women with postnatal depression experience listening visits in primary care? A qualitative interview study. Journal of Reproductive and Infant Psychology. 2006;24:149-62.

Sharp DJ, Chew-Graham C, Tylee A, Lewis G, Howard L, Anderson I, et al. A pragmatic randomised controlled trial to compare antidepressants with a community-based psychosocial intervention for the treatment of women with postnatal depression: the RESPOND trial. Health Technology Assessment. 2010;14:iii-iv, ix-xi, 1-153.

Sherwood RA, Keating J, Kavvadia V, Greenough A, Peters TJ. Substance misuse in early pregnancy and relationship to fetal outcome. European Journal of Peadiatrics. 1999;158:488-92.

Sidebottom AC, Harrison PA, Godecker A, Kim H. Validation of the Patient Health Questionnaire (PHQ)-9 for prenatal depression screening. Archives of Women's Mental Health. 2012;15:367-74.

Silverman K, Svikis D, Eobles E, Stitzer M, Bigelow GE. A reinforcement-based therapeutic workplace for the treatment of drug abuse: six-month abstinence outcomes. Experimental and clinical psychopharmacology. 2001;9:14-23.

Silverman K, Svikis D, Wong CJ, Hampton J, Stitzer ML, Bigelow GE. A reinforcement-based therapeutic workplace for the treatment of drug abuse: three-year abstinence outcomes. Experimental and clinical psychopharmacology. 2002;10:228-240.

Silverstein M, Feinberg E, Cabral H, Sauder S, Egbert L, Schainker E, et al. Problem-solving education to prevent depression among low-income mothers of preterm infants: a randomized controlled pilot trial. Archives of Women's Mental Health. 2011;14:317-24.

Simavli S, Kaygusuz I, Gumus I, Usluogullari B, Yildirim M, Kafali H. Effect of music therapy during vaginal delivery on postpartum pain relief and mental health. Journal of Affective Disorders. 2014;156:194-9.

Simmons RK, Singh G, Maconochie N, Doyle P, Green J. Experience of miscarriage in the UK: qualitative findings from the National Women's Health Study. Social Science and Medicine. 2006;63:1934-46.

Simon GE, Goldberg DP, Von Korff M, Ustun TB. Understanding cross-national differences in depression prevalence. Psychological Medicine. 2002a;32:585-94.

Simon GE, Cunningham ML, Davis RL. Outcomes of prenatal antidepressant exposure. The American Journal of Psychiatry. 2002;159:2055-61.

Sivojelezova A, Shuhaiber S, Sarkissian L, Einarson A, Koren G. Citalopram use in pregnancy: prospective comparative evaluation of pregnancy and fetal outcome. American Journal of Obstetrics and Gynecology. 2005;193:2004-9.

Skapinakis P, Weich S, Lewis G, Singleton N, Araya R. Socio-economic position and common mental disorders: longitudinal study in the general population in the UK. The British Journal of Psychiatry. 2006:189:109-117.

Slade P, Morrell CJ, Rigby A, Ricci K, Spittlehouse J, Brugha TS. Postnatal women's experiences of management of depressive symptoms: a qualitative study. British Journal of General Practice. 2010;60:e440-e448.

Sleed M, Baradon T, Fonagy P. New Beginnings for mothers and babies in prison: a cluster randomized controlled trial. Attachment and Human Development. 2013;15:349-67.

Small R, Lumley J, Donohue L, Potter A, Waldenström U. Randomised controlled trial of midwife led debriefing to reduce maternal depression after operative childbirth. BMJ. 2000;321:1043-7.

Small R, Lumley J, Toomey L. Midwife-led debriefing after operative birth: four to six year follow-up of a randomised trial. BMC Medicine. 2006;4:3.

Smith B, Testa M. The risk of subsequent maltreatment allegations in families with substance-exposed infants. Child Abuse and Neglect. 2002;26:97-114.

Smith EJ, Lui S, Terplan M. Pharmacologic interventions for pregnant women enrolled in alcohol treatment. Cochrane Database of Systematic Reviews. 2009;3:Art. No.:CD007361. DOI: 10.1002/14651858.CD007361.pub2

Smith L, Gibb S. Postnatal support for drug users: evaluation of a specialist health visiting service. Community Practitioner. 2007;80:24-29.

Smith MV, Gotman N, Lin H, Yonkers KA. Do the PHQ-8 and the PHQ-2 accurately screen for depressive disorders in a sample of pregnant women? General Hospital Psychiatry. 2010;32(5):544-8.

Snowdon C, Elbourne D, Forsey M, Alfirevic Z. Information-hungry and disempowered: a qualitative study of women and their partners' experiences of severe postpartum haemorrhage. Midwifery. 2012;28:791-9.

Solmi F, Sallis H, Stahl D, Treasure J, Micali N. Low birth weight in the offspring of women with anorexia nervosa. Epidemiologic Reviews. 2014;36:49-56.

Songoygard KM, Stafne SN, Evensen KA, Salvesen KA, Vik T, Morkved, S. Does exercise during pregnancy prevent postnatal depression? A randomized controlled trial. Acta Obstetricia et Gynecologica Scandinavica. 2012;91:62-7.

Spencer-Smith MM, Spittle AJ, Doyle LW, Lee KJ, Lorefice L, Suetin A, et al. Long-term benefits of home-based preventive care for preterm infants: a randomized trial. Pediatrics. 2012;130:1094-101.

Spies G, Stein DJ, Roos A, Faure SC, Mostert J, Seedat S. Validity of the Kessler 10 (K-10) in detecting DSM-IV defined mood and anxiety disorders among pregnant women. Archives of Women's Mental Health. 2010;69-74:69-74

Spinelli MG, Endicott J. Controlled clinical trial of interpersonal psychotherapy versus parenting education program for depressed pregnant women. American Journal of Psychiatry. 2003;160:555-62.

Spittle AJ, Ferretti C, Anderson PJ, Orton J, Eeles A, Bates L, et al. Improving the outcome of infants born at <30 weeks' gestation - a randomized controlled trial of preventative care at home. BMC Pediatrics. 2009;9:73.

Spittle AJ, Anderson PJ, Lee KJ, Ferretti C, Eeles A, Orton J, et al. Preventative care at home for very preterm infants improves infant and caregiver outcomes at 2 years. Pediatrics. 2010;126:e171-8.

Spitzer RL, Williams JB, Kroenke K, Linzer M, deGruy FV 3rd, Hahn SR, et al. Utility of a new procedure for diagnosing mental disorders in primary care. The PRIME-MD 1000 study. JAMA. 1994;14;272:1749-56.

Spitzer RL, Kroenke K, Williams JBW, for the Patient Health Questionnaire Primary Care Study Group. Validation and utility of a self-report version of PRIME-MD: the PHQ Primary Care Study. JAMA. 1999;282:1737-44.

Springett A, Budd J, Draper ES, Fitzsimons K, Kurinczuk J, Rankin J, et al. Congenital Anomaly Statistics 2011: England and Wales. London: British Isles Network of Congenital Anomaly Registers; 2013.

Stade B, Ali A, Bennett D, Campbell D, Johnston M, Lens C, Tran S, Koren G. The burden of prenatal exposure to alcohol: revised measurement of cost. Canadian Journal of Clinical Pharmacology. 2009a;16:e91-102.

Stade BC, Bailey C, Dzendoletas D, Sgro M, Dowswell T, Bennett D. Psychological and/or educational interventions for reducing alcohol consumption in pregnant women and women planning pregnancy. Cochrane Database of Systematic Reviews. 2009b;2:Art. No.: CD004228. DOI: 10.1002/14651858.CD004228.pub2

Stamp GE, Williams AS, Crowther CA. Evaluation of antenatal and postnatal support to overcome postnatal depression: a randomized, controlled trial. Birth. 1995;22:138-43.

Stanley N, Borthwick R, Macleod A. Antenatal depression: mothers' awareness and professional responses. Primary Health Care Research and Development. 2006;7:257-68.

Stapleton H, Fielder A, Kirkham M. Breast or bottle? eating disordered childbearing women and infant-feeding decisions. Maternal and Child Nutrition. 2008;4:106-20.

Steegers-Theunissen RPM, Renier WO, Borm GF, Thomas CMG, Merkus HMWM, Op De Coul DAW, et al. Factors influencing the risk of abnormal pregnancy outcome in epileptic women: A multi-centre prospective study. Epilepsy Research. 1994;18:261-69.

Stein A, Gath DH, Bucher J, Bond A, Day A, Cooper PJ. The relationship between post-natal depression and mother-child interaction.British Journal of Psychiatry. 1991;158:46-52.

Stein A, Murray L, Cooper P, Fairburn CG. Infant growth in the context of maternal eating disorders and maternal depression: a comparative study. Psychological Medicine. 1996;26:569-74.

Stein A, Woolley H, Senior R, Hertzmann L, Lovel M, Lee J, et al. Treating disturbances in the relationship between mothers with bulimic eating disorders and their infants: a randomized, controlled trial of video feedback. American Journal of Psychiatry. 2006;163:899-906.

Stevenson MD, Scope A, Sutcliffe PA, Booth A, Slade P, Parry G, et al. Group cognitive behavioural therapy for postnatal depression: a systematic review of clinical effectiveness, cost-effectiveness and value of information analyses. Health Technology Assessment. 2010a;14:1-152.

Stevenson MD, Scope A, Sutcliffe PA. The cost-effectiveness of group cognitive behavioral therapy compared with routine primary care for women with postnatal depression in the UK. Value in Health. 2010b;13:580-4.

Stewart RC, Umar E, Tomenson B, Creed F. Validation of screening tools for antenatal depression in Malawi-A comparison of the Edinburgh Postnatal Depression Scale and Self Reporting Questionnaire. Journal of Affective Disorders. 2013;150:1041-7.

Stock A, Chin L, Babl FE, Bevan CA, Donath S, Jordan B. Postnatal depression in mothers bringing infants to the emergency department. Archives of Disease in Childhood. 2013;98:36-40.

Stoesz B, Montgomery J, Smart S, Hellsten LAM. Review of five instruments for the assessment of Asperger's disorder in adults. The Clinical Neuropsychologist. 2011;25:376-401.

Storksen HT, Eberhard-Gran M, Garthus-Niegel S, Eskild A. Fear of childbirth; the relation to anxiety and depression. Acta Obstetricia et Gynecologica Scandinavica. 2011;91:237-42.

Su KP, Huang SY, Chiu TH, Huang KC, Huang CL, Chang HC. Omega-3 fatty acids for major depressive disorder during pregnancy: Results from a randomized, double-blind, placebo-controlled trial. Journal of Clinical Psychiatry. 2008;69:644-51.

Suchman NE, DeCoste C, Leigh D, Borelli J. Reflective functioning in mothers with drug use disorders: implications for dyadic interactions with infants and toddlers. Attachment and Human Development. 2010;12:567-85.

Suchman NE, Decoste C, McMahon TJ, Rounsaville B, Mayes L. The mothers and toddlers program, an attachment-based parenting intervention for substance-using women: results at 6-week follow-up in a randomized clinical pilot. Infant Mental Health Journal. 2011;32:427-49.

Suchman NE, Decoste C, Rosenberger P, McMahon TJ. Attachment-based intervention for substance-using mothers: a preliminary test of the proposed mechanisms of change. Infant Mental Health Journal. 2012;33:360-71.

Suri R, Altshuler L, Hellemann G, Burt VK, Aquino A, Mintz J. Effects of antenatal depression and antidepressant treatment on gestational age at birth and risk of preterm birth. American Journal of Psychiatry. 2007;164:1206-13.

Surkan PJ, Rådestad I, Cnattingius S, Steineck G, Dickman PW. Events after stillbirth in relation to maternal depressive symptoms: a brief report. Birth. 2008;35:153-7.

Sutter-Dallay AL, Murray L, Dequae-Merchadou L, Glatigny-Dallay E, Bourgeois ML, Verdoux H. A prospective longitudinal study of the impact of early postnatal vs. Chronic maternal depressive symptoms on child development. European Journal of Psychiatry. 2011;26:484-9.

Svikis D, Lee JH, Haug N, Stitzer M. Attendance incentives for outpatient treatment: effects in methadone- and non methadone-maintained pregnant drug dependent women. Drug and Alcohol Dependence. 1997;48:33-41.

Swanson KM, Chen H-T, Graham JC, Wojnar DM, Petra A. Resolution of depression and grief during the first year after miscarriage: a randomized controlled clinical trial of couples-focused interventions. Journal of Women's Health. 2009;18:1245-57.

Szumilas M. Explaning Odds Ratios. Journal of Canadian Child and Adolescent Psychiatry. 2010;19:227-229.

Tamaki A. Effectiveness of home visits by mental health nurses for Japanese women with post-partum depression. International Journal of Mental Health Nursing. 2008;17:419-27.

Tandon SD, Cluxton-Keller F, Leis J, Le HN, Perry DF.A comparison of three screening tools to identify perinatal depression among low-income African American women.Journal of Affective Disorders. 2012;136:155-62.

Tandon SD, Perry DF, Mendelson T, Kemp K, Leis JA. Preventing perinatal depression among low-income home visiting clients. Journal of Consulting and Clinical Psychology. 2011;79:707-12.

Tandon SD, Leis JA, Mendelson T, Perry DF, Kemp K. Six-month outcomes from a randomized controlled trial to prevent perinatal depression in low-income home visiting clients. Maternal and Child Health Journal. 2014;18:873-81.

Taylor-Robinson D, Agarwal U, Diggle PJ, Platt MJ, Yoxall B, et al. Quantifying the impact of deprivation on preterm births: a retrospective cohort study. PLoS ONE. 2011;6:8.

Templeton L, Velleman R, Persaud A, Milner P. The experiences of postnatal depression in women from black and minority ethnic communities in Wiltshire, UK. Ethnicity and Health. 2003;8:207-21.

Teng HW, Hsu CS, Shih SM, Lu ML, Pan JJ, Shen WW.Screening postpartum depression with the Taiwanese version of the Edinburgh Postnatal Depression Scale.Comprehensive Psychiatry. 2005;46:261-5.

Terp IM, Mortensen PB. Post-partum psychoses. Clinical diagnoses and relative risk of admission after parturition. The British Journal of Psychiatry. 1998;172:521-6.

Terplan M, Lui S. Psychosocial interventions for pregnant women in outpatient illicit drug treatment programs compared to other interventions. Cochrane Database of Systematic Reviews; 2007;4:Art. No: CD006037. DOI: 10.1002/14651858.CD006037.pub2

Thiagayson P, Krishnaswamy G, Lim ML, Sung SC, Haley CL, Fung DS, et al. Depression and anxiety in Singaporean high-risk pregnancies – prevalence and screening. General Hospital Psychiatry. 2013;35:112-6.

Thomas SV, Ajaykumar B, Sindhu K, Nair MKC, George B, Sarma PS. Motor and mental development of infants exposed to antiepileptic drugs in utero. Epilepsy and Behavior. 2008;13:229-36.

Thompson C, Ostler K, Peveler RC, Baker N, Kinmonth AL. Dimensional perspective on the recognition of depressive symptoms in primary care: The Hampshire Depression Project 3. The British Journal of Psychiatry. 2001:179;317-23.

Thomson G, Downe S. Widening the trauma discourse: the link between childbirth and experiences of abuse. Journal of Psychosomatic Obstetrics and Gynecology. 2008;29:268-73.

Thomson G, Downe S. A hero's tale of childbirth. Midwifery. 2013;29:765-71.

Thurtle V. First time mothers' perceptions of motherhood and PND. CommunityPractitioner. 2003;76:261-5.

Timpano KR, Abramowitz JS, Mahaffey BL, Mitchell MA, Schmidt NB. Efficacy of a prevention program for postpartum obsessive-compulsive symptoms. Journal of Psychiatric Research. 2011;45:1511-7.

Toreki A, Andó B, Keresztúri A, Sikovanyecz J, Dudas RB, Janka Z. The Edinburgh Postnatal Depression Scale: translation and antepartum validation for a Hungarian sample. Midwifery. 2013;29:308-15.

Tran TD, Tran T, La B, Lee D, Rosenthal D, Fisher J. Screening for common perinatal mental disorders in women in the north of Vietnam: a comparison of three psychometric instruments. Asian Journal of Psychiatry. 2011;4:S83-S84.

Truax CB, Carkhuff RR. Toward Effective Counseling and Psychotherapy. Chicago: Aldine; 1967.

Tsartsara E, Johnson MP Women's experience of care at a specialised miscarriage unit: an interpretive phenomenological study. Clinical Effectiveness in Nursing. 2002;6:55-65.

Tseng Y-F, Chen C-H, Lee CS. Effects of listening to music on postpartum stress and anxiety levels. Journal of Clinical Nursing. 2010;19:1049-55.

Turnbull C, Osborn DA. Home visits during pregnancy and after birth for women with an alcohol or drug problem. Cochrane Database of Systematic Reviews. 2012;1:Art. No: CD004456. DOI: 10.1002/14651858.CD004456.pub3

Turner KM, Sharp D, Folkes L, Chew-Graham C. Women's views and experiences of antidepressants as a treatment for postnatal depression: a qualitative study. Family Practice. 2008;25:450-55.

Turner KM, Chew-Graham C, Folkes L, Sharp D. Women's experiences of health visitor delivered listening visits as a treatment for postnatal depression: a qualitative study. Patient Education and Counseling. 2010;78:234-39.Turton P, Hughes P, Evans CD, Fainman D. Incidence, correlates and predictors of post-traumatic stress disorder in the pregnancy after stillbirth. The British Journal of Psychiatry. 2001;178:556-60.

Turton P, Evans C, Hughes P. Long-term psychosocial sequelae of stillbirth: phase II of a nested case-control cohort study. Archives of Womens Mental Health. 2009;12:35-41.

Unger A, Jagsch R, Bäwert A, Winklbaur B, Rohrmeister K, Martin PR, et al. Are male neonates more vulnerable to neonatal abstinence syndrome than female neonates? Gender Medicine. 2011;8:355-64.

Urizar Jr. GG, Muñoz RF. Impact of a prenatal cognitive-behavioral stress management intervention on salivary cortisol levels in low-income mothers and their infants. Psychoneuroendocrinology. 2011;36:1480-94.

Üstün TB, Sartorius N, eds. Mental Illness in General Health Care: An International Study. Chichester: Wiley; 1995.

Uwakwe R, Okonkwo JE. Affective (depressive) morbidity in puerperal Nigerian women: validation of the Edinburgh postnatal depression scale. ActaPsychiatricaScandinavica. 2003;107:251-9.

Vajda FJ, O'Brien TJ, Hitchcock A, Graham J, Cook M, Lander C et al. Critical relationship between sodium valproate dose and human teratogenicity: results of the Australian register of anti-epileptic drugs in pregnancy. Journal of Clinical Neuroscience. 2004;11:854-8.

Vajda FJE, Hitchcock A, Graham J, O'Brien T, Lander C, Eadie M. The Australian Register of Antiepileptic Drugs in Pregnancy: the first 1002 pregnancies. Australian and New Zealand Journal of Obstetrics and Gynaecology. 2007;47:468-74.

Valdimarsdóttir U, Hultman CM, Harlow B, Cnattingius S, Sparén P. Psychotic illness in first-time mothers with no previous psychiatric hospitalizations: a population-based study. PLoS Medicine. 2009;6:e13.

van Doesum KTM, Riksen-Walraven JM, Hosman CMH, Hoefnagels C. A randomized controlled trial of a home-visiting intervention aimed at preventing relationship problems in depressed mothers and their infants. Child Development. 2008;79:547-61.

Veiby G, Daltveit AK, Schjolberg S, Stoltenberg C, Oyen AS, Vollset SE, et al. Exposure to antiepileptic drugs in utero and child development: a prospective population-based study. Epilepsia. 2013;54:1462-72.

Venkatesh KK, Zlotnick C, Triche EW, Ware C, Phipps MG. Accuracy of brief screening tools for identifying postpatrum depression among adolescent mothers. Pediatrics.2014;133:e45-53

Vesga-López O, Blanco C, Keyes K, Olfson M, Grant BF, Hasin DS. Psychiatric disorders in pregnant and postpartum women in the United States. Archive of General Psychiatry. 2008;65:805-15.

Vieten C, Astin J. Effects of a mindfulness-based intervention during pregnancy on prenatal stress and mood: results of a pilot study. Archives of Womens Mental Health. 2008;11:67-74.

Vigod SN, Seeman MV, Ray JG, Anderson GM, Dennis CL, Grigoriadis S, et al. Temporal trends in general and age-specific fertility rates among women with schizophrenia (1996-2009): a population-based study in Ontario, Canada. Schizophrenia Research. 2012;139:169-75.

Vigod SN, Dennis CL, Kurdyak PA, Cairney J, Guttmann A, Taylor VH. Fertility rate trends among adolescent girls with major mental illness: a population-based study. Paediatrics. 2014;133:585-91.Viguera AC, Nonacs R, Cohen LS, Tondo L, Murray A, Baldessarini RJ. Risk of recurrence of bipolar disorder in pregnant and nonpregnant women after discontinuing lithium maintenance. American Journal of Psychiatry. 2000;157:179-84

Viguera AC, Whitfield T, Baldessarini RJ, Newport DJ, Stowe Z, Reminick A, et al. Risk of recurrence in women with bipolar disorder during pregnancy: prospective study of mood stabilizer discontinuation. American Journal of Psychiatry. 2007;164: 1817-24.

Vinten J, Adab N, Kini U, Gorry J, Gregg J, Baker GA. Neuropsychological effects of exposure to anticonvulsant medication in utero. Neurology. 2005;64:949-54.

Waldenström U, Hildingsson I, Ryding EL. Antenatal fear of childbirth and its association with subsequent caesarean section and experience of childbirth. BJOG: An International Journal of Obstetrics and Gynaecology. 2006;113:638-46.

Wan MW, Warren K, Salmon MP, Abel KM. Patterns of maternal responding in postpartum mothers with schizophrenia. Infant Behaviour and Development. 2008;31:532-8.

Wang LH, Lin CC, Chen YH, Lin HC. Increased risk of adverse pregnancy outcomes in women receiving zolpidem during pregnancy. Clinical Pharmacology and Therapeutics. 2010;88:369-74.

Watson HJ, Von Holle A, Hamer RM, Knoph Berg C, Torgersen L, Magnus P, et al. Remission, continuation and incidence of eating disorders during early pregnancy: a validation study in a population-based birth cohort. Psychological Medicine. 2013;43:1723-34.

Webb R, Abel K, Pickles A, Appleby L. Mortality in offspring of parents with psychotic disorders: a critical review and meta-analysis. The American Journal of Psychiatry. 2005;162:1045-56.

Webster J, Linnane J, Roberts J, Starrenburg S, Hinson J, Dibley L. IDentify, Educate and Alert (IDEA) trial: an intervention to reduce postnatal depression. BJOG. 2003;110:842-6.

Weidner K, Bittner A, Junge-Hoffmeister J, Zimmerman K, Siedentopf F, Richter J, et al. A psychosomatic intervention in pregnant in-patient women with prenatal somatic risks. Journal of psychosomatic obstetrics and gynaecology. 2010;31:188-98.

Weissman MM, Markowitz JC, Klerman GL. Comprehensive Guide to Interpersonal Psychotherapy. New York: Basic Books; 2000.

Wen SW, Yang Q, Garner P, Fraser W, Olatunbosun O, Nimrod C, et al. Selective serotonin reuptake inhibitors and adverse pregnancy outcomes. American Journal of Obstetrics and Gynecology. 2006;194:961-6.

Wenzel A, Haugen EN, Jackson LC, Robinson K. Prevalence of generalized anxiety at eight weeks postpartum. Archives of Women's Mental Health. 2003;6:43-9.

Werler MM, Ahrens KA, Bosco JLF, Mitchell AA, Anderka MT, Gilboa SM, et al. Use of Antiepileptic Medications in Pregnancy in Relation to Risks of Birth Defects. Annals of Epidemiology. 2011;21:842-50.

Werrett J, Clifford C. Validation of the Punjabi version of the Edinburgh postnatal depression scale (EPDS). International Journal of Nursing Studies. 2006;43:227-36.

Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Leeflang MM, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Annals of Internal Medicine. 2011;155:529-36.

Whooley MA, Avins AL, Miranda J, Browner WS. Case-finding instruments for depression. Two questions are as good as many. Journal of General Internal Medicine. 1997;12:439-45.

Wichman CL, Moore KM, Lang TR, St Sauver JL, Heise Jr RH, Watson WJ. Congenital heart disease associated with selective serotonin reuptake inhibitor use during pregnancy. Mayo Clinic proceedings. 2009;84:23-7.

Wickberg B, Hwang CP. Counselling of postnatal depression: a controlled study on a population-based Swedish sample. Journal of Affective Disorders. 1996;39:209-16.

Wickberg B, Hwang CP. The Edinburgh Postnatal Depression Scale: validation on a Swedish community sample. ActaPsychiatricaScandinavica. 1996;94:181-4.

Wiggins M, Oakley A, Roberts I, Turner H, Rajan L, Austerberry H, et al. Postnatal support for mothers living in disadvantaged inner city areas: a randomised control trial. Journal of Epidemiology and Community Health. 2005;59:288-95.

Wiklund I, Mohlkert P, Edman G. Evaluation of a brief cognitive intervention in patients with signs of postnatal depression: a randomized controlled trial. Acta Obstetricia et Gynecologica Scandinavica.2010;89:1100-4.

Wikner BN, Stiller CO, Bergman U, Asker C, Kallen B. Use of benzodiazepines and benzodiazepine receptor agonists during pregnancy: neonatal outcome and congenital malformations. Pharmacoepidemiology and Drug Safety. 2007;16:1203-10.

Williams JW Jr, Kerber CA, Mulrow CD, Medina A, Aguilar C. Depressive disorders in primary care: prevalence, functional disability, and identification. Journal of General Internal Medicine. 1995:10;7-12.

Williamson S, Jackson L, Skeoch C, Azzim G,Anderson R.Determination of the prevalence of drug misuse by meconium analysis. Archives of Disease in Childhood. Fetal and Neonatal Edition. 2006;91.

Wilson KL, Zelig CM, Harvey JP, Cunningham BS, Dolinsky BM, Napolitano PG. Persistent pulmonary hypertension of the newborn is associated with mode of delivery and not with maternal use of selective serotonin reuptake inhibitors. American Journal of Perinatology. 2011;28:19-24.

Winhusen T, Kropp F, Babcock D, Hague D, Erickson SJ, Renz C, et al. Motivational enhancement therapy to improve treatment utilization and outcome in pregnant substance users. Journal of Substance Abuse Treatment. 2008;35:161-173.

Winklbaur-Hausknost B, Jagsch R, Graf-Rohrmeister K, Unger A, Baewert A, Langer M, et al. Lessons learned from a comparison of evidence-based research in pregnant opioid-dependent women. Human Psychopharmacology. 2013;28:15-24.

Wirz-Justice A, Bader A, Frisch U, Stieglitz RD, Alder J, Bitzer J, et al. A Randomised, Double-Blind, Placebo-Controlled Study of Light Therapy for Antepartum Depression. Focus on Women's Mental Health. 2011;72:986-993.

Wisner KL, Wheeler SB. Prevention of recurrent postpartum major depression. Hospital and Community Psychiatry. 1994;45:1191-6.

Wisner KL. Antidepressant use and preterm birth: reply. The American Journal of Psychiatry. 2009;166:1189-90.

Wisner KL, Perel JM, Peindl KS, Hanusa BH, Findling RL. Rapport D. Prevention of recurrent postpartum depression: a randomized clinical trial. Journal of Clinical Psychiatry. 2001;62:82-86.

Wisner KL, Perel JM, Peindl KS, Hanusa BH. Timing of depression recurrence in the first year after birth. Journal of Affective Disorders. 2004a;78:249-52.

Wisner KL, Perel JM, Peindl KS, Hanusa BH, Piontek CM et al. Prevention of postpartum depression: a pilot randomized clinical trial. The American Journal of Psychiatry. 2004b;161:1290-92.

Wisner KL, Hunusa BH, Perel JM, Peindl KS, Piontek CM, Sit DKY et al. Postpartum depression: a randomised trial of sertraline versus nortriptyline. Journal of Clinical Psychopharmacology. 2006;26:353-60.

Wisner KL, Sit DK, McShea MC, Rizzo DM, Zoretich RA, Hughes CL, et al. Onset timing, thoughts of self-harm, and diagnoses in postpartum women with screen-positive depression findings. JAMA Psychiatry. 2013;70:490-8.

Wittkowski A, Zumla A, Glendenning S, Fox JRE. The experience of postnatal depression in South Asian mothers living in Great Britain: a qualitative study. Journal of Reproductive and Infant Psychology. 2011;29:480-92.

Wogelius P, Norgaard M, Gislum M, Pedersen L, Munk E, Mortensen PB, et al. Maternal use of selective serotonin reuptake inhibitors and risk of congenital malformations. Epidemiology. 2006;17:701-4.

World Health Organization (WHO). The ICD-10 Classification of Mental and Behavioural Disorders: Clinical Descriptions and Diagnostic Guidelines. Geneva: WHO; 1992.

Wurst KE, Poole C, Ephross SA, Olshan AF. First trimester paroxetine use and the prevalence of congenital, specifically cardiac, defects: A meta-analysis of epidemiological studies. Birth Defects Research Part A - Clinical and Molecular Teratology. 2010:88;159-70.

Yonkers KA, Lin H, Howell HB, Heath AC, Cohen LS. Pharmacologic treatment of postpartum women with new-onset major depressive disorder: A randomized controlled trial with paroxetine. Journal of Clinical Psychiatry. 2008;69:659-65.

Yonkers KA, Forray A, Howell HB, Gotman N, Kershaw T, Rounsaville BJ, et al. Motivational enhancement therapy coupled with cognitive behavioural therapy versus brief advice: a randomized trial for treatment of hazardous substance use in pregnancy and after delivery. General Hospital Psychiatry. 2012;34:439-49.

Yoshida K, Yamashita H, Ueda M, Tashiro N. Postnatal depression in Japanese mothers and the reconsideration of 'Satogaeribunben'. Pediatrics International. 2001;43:189-93.

Zamora J, Abraira V, Muriel A, Khan KS, Coomarasamy A. Meta-DiSc: a software for meta-analysis of test accuracy data. BMC Medical Research Methodology. 2006;6:31.

Zelkowitz P, Milet TH. Screening for post-partum depression in a community sample. Canadian Journal of Psychiatry. 1995;40:80-6.

Zelkowitz P, Feeley N, Shrier I, Stremler R, Westreich R, Dunkley D, et al. The cues and care trial: a randomized controlled trial of an intervention to reduce maternal anxiety and improve developmental outcomes in very low birthweight infants. Neonatal Intensive Care. 2008;22:31-6.

Zelkowitz P, Feeley N, Shrier I, Stremler R, Westreich R, Dunkley D, et al. The cues and care randomized controlled trial of a neonatal intensive care unit intervention: effects on maternal psychological distress and mother-infant interaction. Journal of Developmental and Behavioral Pediatrics. 2011;32:591-9.

Zlotnick C, Johnson SL, Miller IW, Pearlstein T, Howard M. Postpartum depression in women receiving public assistance: pilot study of an interpersonal-therapy-oriented group intervention. American Journal of Psychiatry. 2001;158:638-40.

Zlotnick C, Miller IW, Pearlstein T, Howard M, Sweeney P. A preventive intervention for pregnant women on public assistance at risk for postpartum depression. American Journal of Psychiatry. 2006;163:1443-45.

Zlotnick C, Capezza NM, Parker D. An interpersonally based intervention for low-income pregnant women with intimate partner violence: a pilot study. Archives of Women's Mental Health. 2011;14:55-65.

12 ABBREVIATIONS

ADHD	attention deficit hyperactivity disorder
AD-SUS	Adult Service Use Schedule
AGREE	Appraisal of Guidelines for Research and Evaluation
ALSPAC	*Avon Longitudinal Study of Parents and Children*
ASQ:SE	Ages and Stages Questionnaire: Social-Emotional
AUC	area under the curve
AUDIT	Alcohol Use Disorders Identification Test
BAI	Beck Anxiety Inventory
BDI	Beck Depression Inventory
BSI	Brief Symptom Inventory
CAMHS	child and adolescent mental health services
CBCL	Child Behaviour Checklist
CBT	cognitive behavioural therapy
CBZ	carbamazepine
CCEI	Crown-Crisp Experiential Index
CDSR	Cochrane Database of Systematic Reviews
CENTRAL	Cochrane Database of RCTs and other controlled trials
CES-D	Center for Epidemiologic Studies – Depression scale
CGI	Clinical Global Impressions
CI	confidence interval
CIDI	Composite International Diagnosis Interview
CINAHL	Cumulative Index to Nursing and Allied Health Literature
CIS-R	Clinical Interview Schedule – Revised
CORE-OM	Clinical Outcomes in Routine Evaluation-Outcome Measure
CPN	community psychiatric nurse
DARE	Cochrane Database of Abstracts of Reviews of Effects
DASS-21	Depression, Anxiety, and Stress Scale - 21 items
DHA	docosahexaenoic acid
DSM	*Diagnostic and Statistical Manual of Mental Disorders*
EAS	Emotional Availability Scales
ECT	electroconvulsive therapy
Embase	Excerpta Medica Database
EPA	eicosapentaenoic acid
EPDS	Edinburgh Postnatal Depression Scale
EQ-5D	European Quality of Life – 5 Dimensions

GAD	generalised anxiety disorder
GAD-2, -7	Generalized Anxiety Disorder scale 2 items, 7 items
GDG	Guideline Development Group
GHQ	General Health Questionnaire
GP	general practitioner
GRADE	Grading of Recommendations Assessment, Development and Evaluation
HADS	Hospital Anxiety and Depression Scale
HCHS	Hospital and Community Health Services
HIV	human immunodeficiency virus
HMIC	Health Management Information Consortium
HRQoL	health-related quality of life
HRSD	Hamilton Rating Scale for Depression
HSCL	Hopkins Symptom Checklist
HTA	Health Technology Assessment
IAPT	Improving Access to Psychological Therapies
ICD-10	*International Classification of Diseases*, 10th Revision
ICEP	Infant and Caregiver Engagement Phases
ICER	incremental cost-effectiveness ratio
IES (-R)	Impact of Events Scale (- Revised)
IQ	intelligence quotient
IPT	interpersonal therapy
ISEL	Interpersonal Support Evaluation List
ITT	intention to treat
KID-SCID	Structured Clinical Interview for Childhood Diagnoses
LIFE	Longitudinal Interval Follow-up Examination
LMG	lamotrigine
LOCF	last observation carried forward
MACH	Maternal and Child Health
MADRS	Montgomery-Åsberg Depression Rating Scale
MBU	mother and baby unit
MCMI-III	Millon Clinical Multiaxial Inventory-III
MDD	major depressive disorder
MEDLINE	Medical Literature Analysis and Retrieval System Online
MINI (-PTSD)	Mini International Neuropsychiatric Interview (-Post-Traumatic Stress Disorder)
MMS	Maternal Mood Screener
MRC	Medical Research Council

N	total number of participants
NCAST	Nursing Child Assessment Satellite Training Scale
NCCMH	National Collaborating Centre for Mental Health
NHS	National Health Service
NHS EED	NHS Economic Evaluation Database
NICE	National Institute for Health and Care Excellence
NICU	neonatal intensive care unit
NR	not reported
NTP	nortriptyline
OCD	obsessive-compulsive disorder
OIS	optimal information size
OR	odds ratio
PANDA	Post and Antenatal Depression Association
PCT	primary care trust
PHQ	Patient Health Questionnaire
PICO	Population, Intervention, Comparison and Outcome
PIR-GAS	Parent-Infant Relationship Global Assessment Scale
PPQ	Perinatal PTSD Questionnaire
PRIME-MD	Primary Care Evaluation of Mental Disorders
PSCS	Parenting Sense of Competence Scale
PSI	Parenting Stress Index
PSS	personal social services
PSS-NICU	Parental Stressor Scale-Neonatal Intensive Care
PsycINFO	Psychological Information Database
PTSD	post-traumatic stress disorder
QALY	quality adjusted life year
QIDS	Quick Inventory of Depressive Symptoms
RCT	randomised controlled trial
RDC	research diagnostic criteria
ROC	receiver operating characteristic
RQ	research question
RR	risk ratio
SADS	Schedule for Affective Disorders and Schizophrenia
SAS	Social Adjustment Scale
SCAN	Schedules for Clinical Assessment in Neuropsychiatry
SCID	Structured Clinical Interview for DSM Disorders
SCL-90	Symptom Checklist-90
SD	standard deviation

SERT	sertraline
SES	Rosenberg Self-Esteem Scale
SF (-12, -36)	Short Form Health Survey (-12 items, -36 items)
SMD	standardised mean difference
(S)NRI	(serotonin-) noradrenaline reuptake inhibitor
SPICE	Setting, Perspective, Intervention, Comparison and Evaluation
SPS	Social Provision Scale
SSRI	selective serotonin reuptake inhibitor
STAI (-S, -T)	State-Trait Anxiety Inventory (- State, Trait)
STSI	Short Temperament Scale for Infants
TAU	treatment as usual
TCA	tricyclic antidepressant
VAS	Visual Analogue Scale
VPA	valproate
WCS	worst case scenario
W-DEQ	Wijma Delivery Expectancy Questionnaire
WHO	World Health Organization
WTP	willingness to pay
YBOCS	Yale-Brown Obsessive Compulsive Scale